

			GenScore version 4.5
Copyright (c) 1993 - 2000	Compugen Ltd.		
Run on:	November 29, 2001, 03:54:34	search time 25.08 seconds	(without alignments)
		378.045 Million cell updates/sec	
Title:	US-09-526-329-39		
Perfect score:	685		
Sequence:	1 MASEASVRLGVPPGRLWIQR . . . . . HGQIDSMEQLVLTVOPERKD 128		
Scoring table:	BLOSUM62		
Gapopen:	10.0	Gapext:	0.5
Searched:	522463 seqs, 74073290 residues		
Total number of hits satisfying chosen parameters:	522463		
Minimum DB seq length:	0		
Maximum DB seq length:	200000000		
Post-processing:	Minimum Match 0%		
	Maximum Match 100%		
	Listing first 45 summaries		
Database :			
1: A_GeneSeq_1101:*			
2: /SIDS2/gcadata/geneseq/geneseq/geneseqp/AAI1980.DAT:*			
3: /SIDS2/gcadata/geneseq/geneseq/geneseqp/AAI1982.DAT:*			
4: /SIDS2/gcadata/geneseq/geneseq/geneseqp/AAI1983.DAT:*			
5: /SIDS2/gcadata/geneseq/geneseq/geneseqp/AAI1984.DAT:*			
6: /SIDS2/gcadata/geneseq/geneseqp/AAI1985.DAT:*			
7: /SIDS2/gcadata/geneseq/geneseqp/AAI1986.DAT:*			
8: /SIDS2/gcadata/geneseq/geneseqp/AAI1987.DAT:*			
9: /SIDS2/gcadata/geneseq/geneseqp/AAI1988.DAT:*			
10: /SIDS2/gcadata/geneseq/geneseqp/AAI1989.DAT:*			
11: /SIDS2/gcadata/geneseq/geneseqp/AAI1990.DAT:*			
12: /SIDS2/gcadata/geneseq/geneseqp/AAI1991.DAT:*			
13: /SIDS2/gcadata/geneseq/geneseqp/AAI1992.DAT:*			
14: /SIDS2/gcadata/geneseq/geneseqp/AAI1993.DAT:*			
15: /SIDS2/gcadata/geneseq/geneseqp/AAI1994.DAT:*			
16: /SIDS2/gcadata/geneseq/geneseqp/AAI1995.DAT:*			
17: /SIDS2/gcadata/geneseq/geneseqp/AAI1996.DAT:*			
18: /SIDS2/gcadata/geneseq/geneseqp/AAI1997.DAT:*			
19: /SIDS2/gcadata/geneseq/geneseqp/AAI1998.DAT:*			
20: /SIDS2/gcadata/geneseq/geneseqp/AAI1999.DAT:*			
21: /SIDS2/gcadata/geneseq/geneseqp/AAI2000.DAT:*			
22: /SIDS2/gcadata/geneseqp/AAI2001.DAT:*			
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.			
RESULT 1			
ID AAB18763	AAB18763 standard; Protein; 128 AA.		
XX			
AC AAB18763;			
XX			
DT 22-JAN-2001 (first entry)			
XX			
DE Amino acid sequence of the human Tcl-1b protein.			
XX			
KW Tcl-1; Tcl-1b; T cell malignancy; chromosome 14 abnormality; lymphoma; T-cell leukaemia; immunodeficiency syndrome; ataxia-teleangiectasia.			
XX			
OS Homo sapiens.			
XX			
PN WO200051569-A1.			
XX			
PD 21-SEP-2000.			
XX			
PF 15-MAR-2000; 2000WO-US06612.			
XX			
PR 15-MAR-1999; 99US-0124714.			
XX			
PA (UYTE-) UNIV JEFFERSON THOMAS.			
XX			
PI Croce CM, Pekarsky Y;			
XX			
DR WPI: 2000-611514/58.			
DR N-PSDB; AAC15822.			
XX			
PT Novel nucleic acid of Tcl-1 gene family, Tcl-1b, expressed in low levels in normal bone marrow and peripheral lymphocytes, but activated in T-cell leukaemia and lymphoma, used to identify chromosome 14 abnormalities -			
PT			
protein encoded by			
protein encoded by			

## ALIGNMENTS

12	69	10.1	1012	16	AAB85439	IBDV GLS structure
13	69	10.1	1012	18	AAM14737	IBDV strain GLS 1a
14	69	10.1	1021	7	AAP1316	Infectious bursal
15	69	10.1	1023	12	AAM15054	Peptidopeptide encode
16	68.5	10.0	544	19	AAB85028	CD22-green fluores
17	68.5	10.0	980	12	AAM1113	Equine C-terminal
18	68	9.9	1012	21	AAY7057	Infectious bursal
19	68	9.9	1012	22	AAB31631	Amino acid sequenc
20	68	9.9	1013	19	AAY54377	IBDV VP2-VP4-VP3 P
21	67.5	9.9	195	21	AAY94880	Human Protein clon
22	67.5	9.9	635	22	AAB92860	C glutamicum prote
23	67	9.8	222	20	AAY97933	Pineapple polyphen
24	67	9.8	310	22	AAB81362	Human ARP protein
25	67	9.7	467	22	AAB94032	Pineapple polyphen
26	67	9.8	619	20	AAY97959	Human Protein sequ
27	67	9.8	684	22	AAB94404	IBDV strain ENCAR
28	67	9.8	1012	13	AAB29277	IBDV structural pr
29	67	9.8	1012	14	AAB4124	Consensus amino ac
30	67	9.8	1012	22	AAB31639	Human Protein kina
31	66.5	9.7	1000	22	AAB38700	Novel Protein sequ
32	66.5	9.7	1490	22	AAB65544	Protein having aur
33	66	9.6	562	20	AAY41268	Arabidopsis thalia
34	66	9.6	987	21	AAG40315	Arabidopsis thalia
35	66	9.6	997	21	AAG44314	Arabidopsis thalia
36	66	9.6	1001	21	AAG40313	Arabidopsis thalia
37	66	9.6	1012	14	AAG38700	IBDV gene product.
38	66	9.6	1012	15	AAR51171	Sequence of wild t
39	66	9.6	1012	15	AAR51049	Sequence of antigen
40	66	9.5	1012	15	AAR53227	Segment A of IBDV
41	66	9.6	1012	19	AAB54379	IBDV polyprotein V
42	66	9.6	1012	20	AAB54375	IBDV VP2-VP3-VP4 P
43	66	9.6	1012	22	AAB31630	Amino acid sequenc
44	65.5	9.6	762	21	AAG30764	Arabidopsis thalia
45	65.5	9.6	764	21	AAG30763	Arabidopsis thalia



CC library. Recombinant TCI-1 can be obt'd. by incorporation of the cDNA into a vector and expression in *Escherichia coli* transfectants. The TCI-1 protein and antibodies raised against it can be used for the diagnosis and treatment of conditions associated with increased expression of TCI-1 proteins and/or with chromosomal abnormalities, esp. T cell leukaemia and lymphomas with chromosome 14 abnormalities.

XX Sequence 113 AA;

Query Match 21.8%; Score 149; DB 17; Length 113; Best Local Similarity 29.2%; Pred. No. 7.6e-10; Matches 33; Conservative 21; Mismatches 45; Indels 14; Gaps 1;

Qy 13 PGRIWIQRPGLYDEBEGRWWTVVRFNPNSREWARASQSGRSYEPSTVHLWOMAVHRE 72  
Db 60 pmtptqigslipimwqlypdgryrssdssfwrlivhikgdedmlellpd 112

RESULT 4

ID AAM39823  
XX AAM39823 standard; Protein; 114 AA.

AC AAM39823;  
XX  
DT 22-OCT-2001 (first entry)

DE Human polypeptide SEQ ID NO 2968.

KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic; chemokinetic; thrombolytic; drug screening; arthritis; inflammation; leukaemia.

XX Homo sapiens.

XN WO200153312-A1.

XX  
DD 26-JUL-2001.

PR 26-DEC-2000; 2000WO-US34263.

XX  
PR 21-JAN-2000; 20000S-0488725.

PR 25-APR-2000; 20000S-0552317.

PR 09-JUL-2000; 20000S-0598042.

PR 19-JUL-2000; 20000S-0520312.

PR 03-AUG-2000; 20000S-0653450.

PR 14-SEP-2000; 20000S-0652191.

PR 19-OCT-2000; 20000S-0653056.

PR 29-NOV-2000; 20000S-0727344.

XX  
PA (HYSE-) HYSEQ INC.

PT Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D; PI Wang J, Wang Z, Wehman T, Xu C, Xue AJ, Yang Y, Zhang J; XX Zhao QA, Zhou P, Goodrich R, Dumanac RT; DR WPI; 2001-442253/47.

PR N-PSDB; RA158979.

PT Novel nucleic acids and polypeptides, useful for treating disorders such as central nervous system injuries -

XX Example 4; SEQ ID NO 2968; 10078P; English.

CC The invention relates to human nucleic acids (AA157798-AA161369) and

CC the encoded polypeptides (AAM38642-AM42213) with nootropic, CC immunosuppressant and cytostatic activity. The polynucleotides are useful CC in gene therapy. A composition containing a polypeptide or polynucleotide CC of the invention may be used to treat diseases of the peripheral nervous CC system, such as peripheral nervous injuries, peripheral neuropathy and CC localised neuropathies and central nervous system diseases, such as CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the CC utilisation of the activities such as: immune system suppression, CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic CC and thrombolytic activity, cancer diagnosis and therapy, drug screening, CC assays for receptor activity, arthritis and inflammation, leukaemias and CC C.N.S disorders.

Note: The sequence data for this patent did not form part of the printed CC specification.

XX Sequence 114 AA;

Query Match 21.8%; Score 149; DB 22; Length 114; Best Local Similarity 29.2%; Pred. No. 7.7e-10; Matches 33; Conservative 21; Mismatches 45; Indels 14; Gaps 1;

Qy 13 PGRIWIQRPGLYDEBEGRWWTVVRFNPNSREWARASQSGRSYEPSTVHLWOMAVHRE 72  
Db 15 pmtptqigslipimwqlypdgryrssdssfwrlivhikgdedmlellpd 60

RESULT 5

ID AAM41609  
XX AAM41609 standard; Protein; 129 AA.

AC AAM41609;  
XX  
DT 22-OCT-2001 (first entry)

DE Human polypeptide SEQ ID NO 6540.

XX  
OS Homo sapiens.

XX  
PR WO200153312-A1.

XX  
PD 26-JUL-2001.

XX  
PR 26-DEC-2000; 2000WO-US34263.

XX  
PR 21-JAN-2000; 20000S-0488725.

XX  
PR 25-APR-2000; 20000S-0552317.

XX  
PR 09-JUL-2000; 20000S-0598042.

XX  
PR 19-JUL-2000; 20000S-0520312.

XX  
PR 03-AUG-2000; 20000S-0653450.

XX  
PR 14-SEP-2000; 20000S-0652191.

XX  
PR 19-OCT-2000; 20000S-0653056.

XX  
PR 29-NOV-2000; 20000S-0727344.

XX  
PA (HYSE-) HYSEQ INC.

XX  
PR 21-JAN-2000; 20000S-0488725.

XX  
PR 25-APR-2000; 20000S-0552317.

XX  
PR 09-JUL-2000; 20000S-0598042.

XX  
PR 19-JUL-2000; 20000S-0520312.

XX  
PR 03-AUG-2000; 20000S-0653450.

XX  
PR 14-SEP-2000; 20000S-0652191.

XX  
PR 19-OCT-2000; 20000S-0653056.

XX  
PR 29-NOV-2000; 20000S-0727344.

XX  
PA (HYSE-) HYSEQ INC.

XX  
PT Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D; PI Wang J, Wang Z, Wehman T, Xu C, Xue AJ, Yang Y, Zhang J; XX Zhao QA, Zhou P, Goodrich R, Dumanac RT; DR WPI; 2001-442253/47.

DR N-PSDB; AAI60765.

XX Novel nucleic acids and polypeptides, useful for treating disorders

PT such as central nervous system injuries -

PS Example 2; SEQ ID NO 6540; 10078pp; English.

XX

CC The invention relates to human nucleic acids (AAI57798-AAI61369) and

CC the encoded polypeptides (AAM3842-AAM2213) with notropic, immunosuppressant and cyrostatic activity. The polynucleotides are useful

CC in gene therapy. A composition containing a polypeptide or polynucleotide

CC of the invention may be used to treat diseases of the peripheral nervous

CC system, such as peripheral neuropathy and

CC localised neuropathies and central nervous system diseases, such as

CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic

CC lateral sclerosis and Sly-Drager Syndrome. Other uses include the

CC utilisation of the activities such as: immune system suppression,

CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic

CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,

CC assays for receptor activity, arthritis and inflammation, leukemias and

CC C.N.S disorders.

CC Note: The sequence data for this patent did not form part of the printed

CC specification.

XX

SQ Sequence 129 AA:

Query Match 21.8%; Score 149; DB 22; Length 129;

Best Local Similarity 29.8%; Pred. No. 9.1e-10; Mismatches 45; Indels 14; Gaps 1;

Matches 33; Conservative 21; Mismatches 45; Indels 14; Gaps 1;

Qy 13 PQRWLRQRCIYDEEGRTWVYVWRNPSRERWAASQSRPSTVHOMAVHIRE 72

Db 30 pdrlwakewkvyldeqkqaxcpliekrdrqlqxvrlrredvigrbpmtpixgpl-- 75

Qy 73 LWSGOMPRSQLPAWQYLPKRRRAAASSFETIAHGGQIDSMEQLTYOPE 125

Db 76 pmpptqdgslpilmwqlypdgryssdssfrvlyhikidgvedmllelpd 128

RESULT 6

ID AAR94975 standard; Protein: 113 AA.

XX

AC AAR94975;

XX

DT 23-AUG-1995 (first entry)

XX

DE Human TCR-1 polypeptide.

XX

OS Homo sapiens.

XX

PH Key Location/Qualifiers

FT Misc-difference 31 /note= "unidentified amino acid"

FT Misc-difference 46 /note= "unidentified amino acid"

FT Misc-difference 47 /note= "unidentified amino acid"

FT Misc-difference 55 /note= "unidentified amino acid"

FT Misc-difference 55 /note= "unidentified amino acid"

XX

W09613514-A1.

XX

PN 09-MAY-1996.

XX

PF 23-OCT-1995; 95WO-US13663.

XX

PR 27-OCT-1994; 94US-0330272.

XX

(RAGG-) RAGGIO-ITALGENE SPA.

PA (UYJE-) UNIV JEFFERSON THOMAS.

XX

PI Croce CM, Russo G;

XX

DR N-PSDB; AAT18877.

XX

PT New TCL-1 protein and gene associated with chromosome 14

PT abnormalities - useful to develop prods. for detection, treatment

PT and prevention of diseases such as T-cell leukaemia(s) and

PT lymphoma(s).

XX

PS Disclosure: Page 69-72; 105pp; English.

XX

CC The amino acid sequence (AAT94975) of human TCR-1 protein was

CC deduced from the exon sequences of the TCR-1 gene (AAT8877).

CC A sequence deduced from a cDNA clone is given in AAR94974.

CC The TCR-1 gene is expressed at high levels in leukemic cells

CC carrying a t(14;14)(q11;q32) translocation or an inv(14)(q11;q32)

CC inversion. The TCR-1 protein, and antibodies raised against it,

CC can be used for the diagnosis or treatment of conditions associated

CC with increased expression of TCR-1 proteins and/or with chromosome

CC 14 abnormalities, esp. T-cell leukaemia and lymphoma.

XX

SQ Sequence 113 AA:

Query Match 19.3%; Score 132; DB 17; Length 113;

Best Local Similarity 26.8%; Pred. No. 7.6e-08; Mismatches 39; Indels 34; Gaps 3;

Matches 33; Conservative 17; Mismatches 39; Indels 34; Gaps 3;

Qy 13 PSLWLRQPGIYDEEGRTWVYVWRNPSRERWAASQSRPSTVHOMAVHIRE 62

Db 14 pdrlwakewkvyldeqkqaxcpliekrdrqlqxvrlrredvigrbpmtpixgpl-- 70

Qy 63 LWMOMAVHIRELLSSGQMFSQLPAWQYLPKRRRAAASSFETIAHGGQIDSMEQLTYOPE 122

Db 71 -----lpmwqlypdgryssdssfrvlyhikidgvedmllelp 109

Qy 123 QPE 125

Db 110 lpd 112

RESULT 7

ID AAE05866 standard; Protein: 513 AA.

XX

AC AAE05866;

XX

DT 24-SEP-2001 (first entry)

XX

DE Pseudomonas stutzeri open reading frame-0 (ORF-0) protein.

XX

PH Open reading frame-0; ORF-0; pyridine-2,6-bis (thiocarboxylate);

KW PdTC; environmental remediation; phytoremediation; biaccumulation;

KW water purification; solution mining; remediation; immobilisation;

KW detoxification; redox state modifier; metal ion reactivity;

KW carbon tetrachloride; metal.

XX

OS Pseudomonas stutzeri.

XX

PN W0200153309-A1.

XX

PD 26-JUL-2001.

XX

PF 19-JAN-2001; 2001WO-US02386.

XX

PR 20-JAN-2000; 2000US-0177251.

XX

PA (IDAH-) IDAHO RES FOUND INC.

PA (PASZL-) PASZCZYNKI A.

PA (SEBA-) SEBAT J L.

XX  
PI Paszczynski A, Sebat JL, Lewis TA, Crawford RL, Cortese MS;  
XX  
DR WPI; 2001-465361/50.  
XX  
DR N-PSDB; AAD11187.  
XX  
PT New isolated nucleic acids, useful for producing enzymes required to  
PT produce pyridine-2,6-bis (thiocarboxylate), especially useful for  
PT reducing the amount of metal or carbon tetrachloride in a substrate,  
e.g. soil or water  
XX  
PS Example 1; Page 164-166; 172pp; English.  
XX  
CC The present sequence is a *Pseudomonas stutzeri* open reading frame-O  
CC (ORF-O) protein. The *Pseudomonas stutzeri* genome includes ORFs that  
CC encode enzymes required for biosynthesis of pyridine-2,6-bis  
CC (thiocarboxylate) (PTC). The ORFs encoding PTC are especially useful  
CC in environmental remediation methods, e.g. phytoremediation,  
CC biaccumulation, water purification, waste water purification, solution  
CC mining, mobilisation, immobilisation, detoxification, redox state  
CC modifier or modification of metal ion reactivity. In particular, the  
CC ORFs are useful for degrading carbon tetrachloride and removing metals  
XX  
Sequence 513 AA;

Query Match 10.7%; Score 73; DB 22; Length 513;  
Best Local Similarity 24.3%; Pred. No. 4.9%;  
Matches 25; Conservative 16; Mismatches 42; Indels 20; Gaps 4;

QY 39 FNPSRREWARASQSR-----VEPS-----ITVHLWQMAVTRELLSSQMP 80  
45 fgprtstwsqskagrlyvhchpprqtcarrrhryqfvschrlhllphkrarrgvp 104  
QY 81 FSLPATAWQWQYPCRKYRAADS-SFWEIADHGQIDSMEQLVT 121  
DB 105 19q1pprwwqldpsqlsrlslagtpandmidfsipnevqmlvst 147

RESULT 8  
AAW71290 standard; Protein: 906 AA.  
XX  
AC AAW71290;  
XX  
DT 04-DEC-1998 (first entry)  
XX  
DE Potato starch branching enzyme.  
XX  
KW Potato; starch branching enzyme; SBE; class A; class B; modified starch;  
XX  
OS Solanum tuberosum.  
XX  
RN WO937214-A1.  
XX  
PD 27-AUG-1998.  
XX  
PF 23-FEB-1998; 98WO-IB00295.  
XX  
PR 24-MAR-1997; 97GB-0006075.  
XX  
PR 21-FEB-1997; 97GB-0003672.  
XX  
PA (DANI-) DANISCO AS.  
XX  
PI Poulsen P;  
XX  
DR WPI; 1998-48001/41.  
XX  
DR N-PSDB; AAV54979.  
PT Modifying starch producing organisms, particularly plants - by  
expression of nucleotide sequences which code for introns of starch

PT branching enzyme sequences  
XX  
PS Disclosure; Fig 8; 108pp; English.  
XX  
CC The present sequence represents a potato starch branching enzyme (SBE);  
CC The specification describes a method for affecting enzymatic activity in  
CC a plant. The method comprises expressing in the plant a nucleotide  
CC sequence encoding, partially or completely, an intron of a class A potato  
CC sequence which codes partially or completely, for an intron of a class B  
CC SBE in a sense or antisense orientation, and where the nucleotide  
CC sequence does not contain a sequence that is sense to an exon sequence  
genetically associated with the intron. The method can be used for providing  
CC improved starches whose properties would satisfy various industrial  
CC requirements. They can be used for preparing tailor-made starches in  
CC plants which could replace the post-harvest modified starches. They can  
also be used for expressing genes of interest, e.g. pharmaceutically  
active proteins or enzymes or proteins which are beneficial to plants.  
XX  
SQ Sequence 906 AA;

Query Match 10.2%; Score 70; DB 19; Length 906;  
Best Local Similarity 24.2%; Pred. No. 2.4;  
Matches 31; Conservative 17; Mismatches 34; Indels 46; Gaps 8;

QY 17 WIORPGIYEDDEGRWTVWRPNPSREWARA-SQSSRVEPSITVHLWQMAVTR-----72  
DB 621 wldfp-----regnw-----syakcrqrgnlaosehlyrk--fmnaidramnsidekf 667  
QY 73 -LISSGOMPFES-----QDPAWQWQYPCRKYRAADS-SFWEIADHGQIDSMEQLVT 103  
DB 668 sflasgqkivssmddkvuvfergdvlvfnfhpkntqeykvqcdlpgkyvalds 727  
QY 104 WEIADHGQ 111  
DB 728 wefqghgr 735

RESULT 9  
AAW69300  
ID AAW69300 standard; Protein: 906 AA.  
XX  
AC AAW69300;  
XX  
DT 07-DEC-1998 (first entry)  
XX  
DE Potato class B starch branching enzyme.  
XX  
KW Starch branching enzyme; SBE; potato; antisense; amylopectin;  
XX  
OS Solanum tuberosum cv. Desiree.  
XX  
RN WO937213-A1.  
XX  
PD 27-AUG-1998.  
XX  
PF 23-FEB-1998; 98WO-IB00270.  
XX  
PR 24-MAR-1997; 97GB-0006050.  
XX  
PR 21-FEB-1997; 97GB-0003663.  
XX  
PA (DANI-) DANISCO AS.  
XX  
PI Poulsen P;  
XX  
DR WPI; 1998-467573/40.  
XX  
DR N-PSDB; AAV55069.  
PT Affecting enzymatic activity using antisense intron inhibition -  
especially of starch branching enzyme in plants, useful to alter



CC The sequences given in AAI14738-41 are encoded by fragments of the  
 CC infectious bursal disease virus (IBDV) large segment from the GLS 5  
 CC strain. These sequences are encoded by clones GLS-1 to GLS-4. Proteins  
 CC such as these, derived from IBDV, esp. VP2 which is the major host  
 CC protective immunogen of IBDV, may be used to prepare vaccines against  
 CC IBDV infection/Gumboro disease.

SQ Sequence 540 AA;

Query Match

Best Local Similarity 23.8%; Score 69; DB 18; Length 540;

Matches 30; Conservative 18; Mismatches 50; Indels 28; Gaps 5;

QY 8 RLGVPPGRILWIRPGIYDEDEGRITWVTVVRFNSRREWAR-----ASQGSRYEP 57

Db 698 rlg---lklagpafdnqgnwafkrifpnnprwdtripylnlpypnraqy- 750

QY 58 SITVHLWQMAVHTREL-LSSGOMPFSQSQLPAWQLYPGRKYRAADSSFWEIAHDGQIDSM 115

Db 751 ---hlanasefkettpelesavramaaasvdpl----fqsalsvfmleengivtdm 801

QY 116 EQVLT 121

Db 802 anfalls 807

RESULT 13

AAW14737

ID AAW14737 standard; Protein; 1012 AA.

XX

AC AAW14737;

XX

DT 08-MAY-1997 (first entry)

DE IBDV strain GLS large segment derived protein.

XX

KW Infectious bursal disease virus; IBDV; large segment; GLS 5 strain;

KW structural protein; VP2; VP3; VP4; immunogen; vaccine;

KW Gumboro disease.

OS Infectious bursal disease virus.

XX

PN US5595912-A.

XX

PD 21-JAN-1997.

XX

PP 04-MAY-1990; 900US-0519202.

XX

PR 23-MAR-1994; 940US-0216276.

XX

PR 04-MAY-1990; 900US-0519202.

XX

PR 28-JUN-1993; 930US-0083784.

XX

PA (UYMA-) UNIV MARYLAND BALTIMORE.

XX

PI Snyder D, Vakharia V;

XX

DR WPI-1997-10784/10.

XX

DR N-FSDB, AAT53055.

XX

PT Infectious bursal disease virus nucleic acids - for prodn. of

PT vaccines for poultry for protection against Gumboro disease

XX

PS Claim 1; Column 43-52; 41pp; English.

XX

CC This sequence is encoded by a fragment of the infectious bursal disease

CC virus (IBDV) large segment from the GLS 5 strain. This sequence

CC represents the structural proteins VP2, VP3 and VP4. These proteins

CC esp. VP2 which is the major host protective immunogen of IBDV, may be

CC used to prepare vaccines against IBDV infection/Gumboro disease.

XX

SQ Sequence 1012 AA;

Query Match

Best Local Similarity 23.8%; Score 69; DB 18; Length 1012;

Matches 30; Conservative 18; Mismatches 50; Indels 28; Gaps 5;

QY 8 RLGVPPGRILWIRPGIYDEDEGRITWVTVVRFNSRREWAR-----ASQGSRYEP 57

Db 698 rlg---lklagpafdnqgnwafkrifpnnprwdtripylnlpypnraqy- 750

QY 58 SITVHLWQMAVHTREL-LSSGOMPFSQSQLPAWQLYPGRKYRAADSSFWEIAHDGQIDSM 115

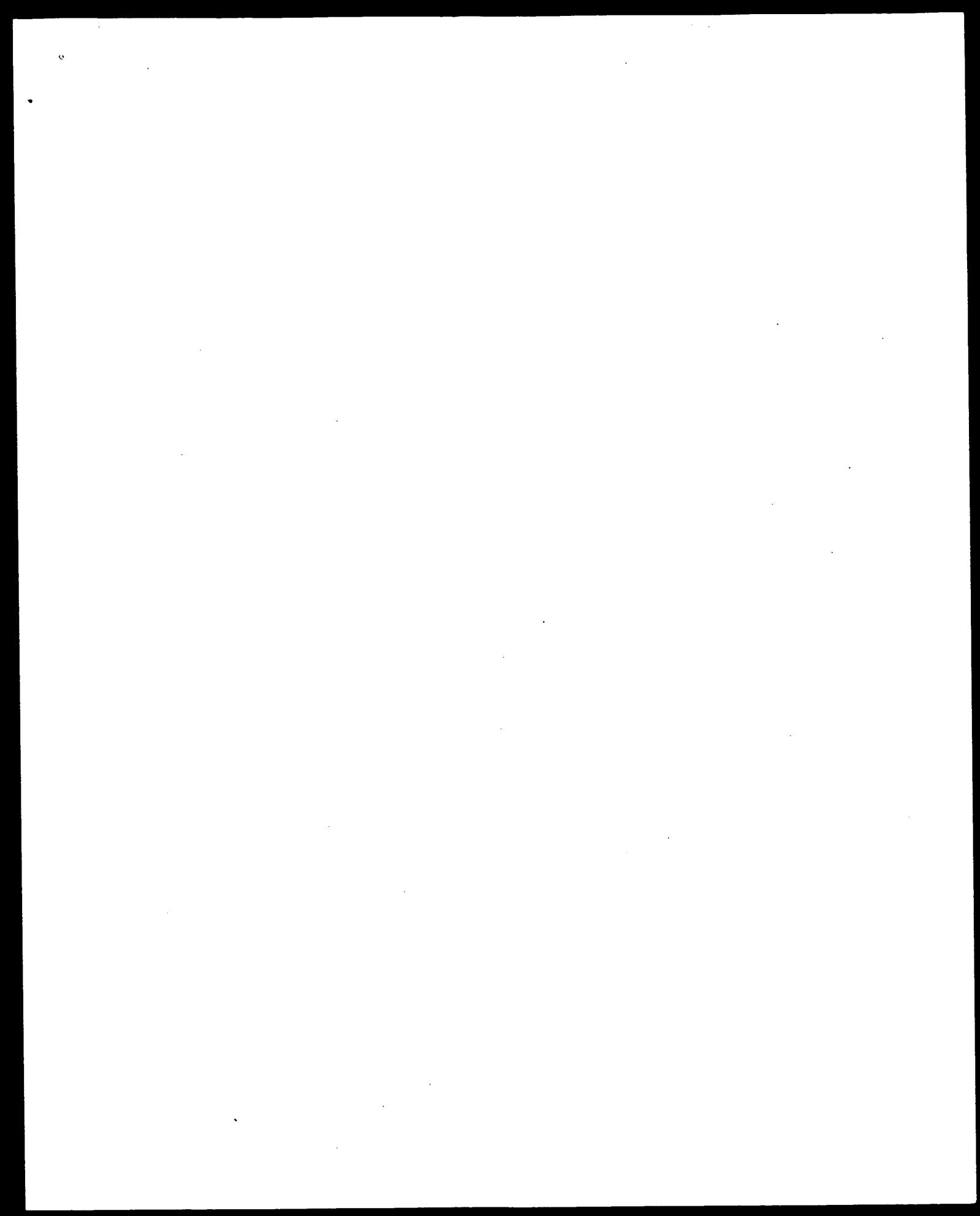


© Fri Nov 30 11:03:09 2001

us-09-526-329-39.rag

Page 9

Search completed: November 29, 2001, 04:03:10  
Job time: 516 sec







; FILING DATE:  
 ; CLASSIFICATION:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: MISROCK, S. Leslie  
 ; REGISTRATION NUMBER: 18,972  
 ; REFERENCE/DOCKET NUMBER: 6754-027  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (212) 790-8864/9741  
 ; TELEFAX: (212) 790-9090  
 ; TELEX: 66141 PENNIE  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 113 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; PCT-US95-13663-2  
 ; US-08-216-276A-33  
 ; Sequence 33, Application US/08216276A  
 ; Patent No. 5535912  
 ; GENERAL INFORMATION:  
 ; APPLICANT: VAKHARIA, VIKRAM  
 ; TITLE OF INVENTION: SPECIFIC DNA AND RNA SEQUENCES  
 ; ASSOCIATED WITH US HBV VARIANTS, VECTOR CARRYING DNA  
 ; TITLE OF INVENTION: SEQUENCES, HOST CARRYING CLONED VECTOR, DEDUCED AMINO ACID  
 ; TITLE OF INVENTION: SEQUENCES, VACCINE AND METHOD OF VACCINATION  
 ; NUMBER OF SEQUENCES: 34  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: OBLON, SPIVAK, McCLELLAND, MATER & NEUSTADT,  
 ; STREET: 1755 S. Jefferson Davis Highway, Suite 400  
 ; CITY: Arlington  
 ; STATE: Virginia  
 ; COUNTRY: U.S.A.  
 ; ZIP: 22202  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patentin Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/216,276A  
 ; FILING DATE: 23-MAR-1994  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/083,784  
 ; FILING DATE: 28-JUN-1993  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 07/519,202  
 ; FILING DATE: 04-MAY-1990  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 07/227,311  
 ; FILING DATE: 02-AUG-1988  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Kelber, Steven B.  
 ; REGISTRATION NUMBER: 30,073  
 ; REFERENCE/DOCKET NUMBER: 2747-054-27 CIP  
 ;  
 ; Query Match 21.8%; Score 149, DB 5; Length 113;  
 ; Best Local Similarity 29.2%; Pred. No. 1.4e-10;  
 ; Matches 33; Conservative 21; Mismatches 45; Indels 14; Gaps 1;  
 ; Qy 13 PRGMLIOPGPIYEDDEGRWVWVVRFNRSRRWAR-----ASQGSRYEP 57  
 ; Db 14 PRDLWAWKEVKYLDEKQHAWLPLTIKDRQ-----LRVLLRDEVVLGR 59  
 ; Qy 73 LLSGGMPFSQLAWQLYPGRKRYADSSFWETADHGQIDSMEOLVLYQPE 125  
 ; Db 60 PMPPTQGSPSLPNTWQLYPGRKYSQSDSFWRWVWHIKDGVEMELDFD 112  
 ;  
 ; Query Match 10.1%; Score 69, DB 1; Length 476;  
 ; Best Local Similarity 23.8%; Pred. No. 3.8; Matches 30; Conservative 18; Mismatches 50; Indels 28; Gaps 5;  
 ; Gaps 30; Matches 30; Conservative 18; Mismatches 50; Indels 28; Gaps 5;  
 ; Qy 8 RLGVPPGRLWITOPGPIYEDDEGRWVWVVRFNRSRRWAR-----ASQGSRYEP 57  
 ; Db 162 RLG---IKLAGPKFDVNTGPNWATKRFPHNPRWDRLPVLNPYLPNPGROY-- 214  
 ; Qy 58 SITVHLMQMVAYHTREL--LSSGMPFSQLAWQLYPGRKRYADSSFWETADHGQIDS 115  
 ; Db 215 ---HLAMAASEFKETPELESAVRAMEAAASVDP-----FQSLASVFWLEENGIVDM 265  
 ; Qy 116 EQVLT 121  
 ; Db 266 ANFALS 271  
 ;  
 ; Query Match 6  
 ; Best Local Similarity 23.8%; Pred. No. 3.8; Matches 31; Conservative 18; Mismatches 50; Indels 28; Gaps 5;  
 ; Gaps 31; Matches 31; Conservative 18; Mismatches 50; Indels 28; Gaps 5;  
 ; Qy 10 US-08-216-276A-31  
 ; Sequence 31, Application US/08216276A  
 ; Patent No. 5535912  
 ; GENERAL INFORMATION:  
 ; APPLICANT: VAKHARIA, VIKRAM  
 ; TITLE OF INVENTION: SPECIFIC DNA AND RNA SEQUENCES  
 ; ASSOCIATED WITH US HBV VARIANTS, VECTOR CARRYING DNA  
 ; TITLE OF INVENTION: SEQUENCES, HOST CARRYING CLONED VECTOR, DEDUCED AMINO ACID  
 ; TITLE OF INVENTION: SEQUENCES, VACCINE AND METHOD OF VACCINATION  
 ; NUMBER OF SEQUENCES: 34  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: OBLON, SPIVAK, McCLELLAND, MATER & NEUSTADT,  
 ; ADDRESSER: P.C.  
 ; STREET: 1755 S. Jefferson Davis Highway, Suite 400  
 ; CITY: Arlington  
 ; STATE: Virginia  
 ; COUNTRY: U.S.A.  
 ; ZIP: 22202  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patentin Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/216,276A  
 ; FILING DATE: 23-MAR-1994  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/083,784  
 ; FILING DATE: 28-JUN-1993  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 07/519,202  
 ; FILING DATE: 04-MAY-1990  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 07/227,311  
 ; FILING DATE: 02-AUG-1988  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Kelber, Steven B.  
 ; REGISTRATION NUMBER: 30,073  
 ; REFERENCE/DOCKET NUMBER: 2747-054-27 CIP



ORGANISM: Infectious bursal disease virus  
STRATN: GLS  
US-08-219-262B-1

Db 751 ---HLAMASEFKETPELESAVRAMEAAASVDPL---FOSALSYWMLEENGIVWDM 801

QY 116 EQLVLP 121  
Dh 802 amper 607  
|:

RESULT  
US-08-219-262B-9  
; Sequence 9, Application US/08219262B

PATENT NO. 5788970  
GENERAL INFORMATION:  
APPLICANT: VAKHARIA, VIKRAM

APPLICANT: SNYDER, DAVID B  
APPLICANT: MENGEI-WEHRSAT, STEPHANIE A  
TITLE OR INVENTION: ~~SYSTEMS FOR DETERMINING~~

INVENTION: CHICKEN INFECTIOUS BURSAL DISEASE VIRUS  
TITLE OF INVENTION: CCNA CLONES, EXPRESSION PRODUCTS AND VACCINES BASED  
TITLE OF INVENTION: THEREON  
INVENTOR: AN INVENTOR

NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT

STREET: 1175 S. JEFFERSON DAVID HIGHWAY, FOURTH FLOOR  
CITY: ALEXANDRIA  
STATE: VIRGINIA

MEDIUM TYPE: FLOPPY disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: DOS

OVERDRAFTING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PATENTIN RELEASE #1.0, Version #1.30  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US08/219,262B  
FILING DATE: 29-MAR-1994  
CLASSIFICATION: 435

ATTORNEY / AGENT INFORMATION:  
NAME: OBLON, NORMAN F  
REGISTRATION NUMBER: 24 619

REDOCK 24, 010  
REFERENCE/DOCKET NUMBER: 2747-047-27  
TELECOMMUNICATION INFORMATION:  
MAY 2004

TELEPHONE: (003) 413-3000  
TELEFAX: (703) 413-2220  
TELEX: 248855 OPAT UR

INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1012 amino acids

TYPE: amino acid  
STRANGENESS: 0  
TOPOLOGY: unknown

MOLECULE TYPE: protein  
ORIGINAL SOURCE: *Opuntia*  
CULTIVAR: *Frijoles*

ORGANISM: INFECTIOUS bursal disease virus  
STRAIN: 002-73  
S-08-219-262B-9

Query Match Score 69; DB 1; Length 1012;  
best Local Similarity 10.1%;  
Pred. No. 11;

Matches	31;	Conservative	16;	Mismatches	51;	Indels	28;	Gaps
Qy	8	RIGVPPGRMIORGIEDEBGRVVVVFNPNSRREWAA						
Db	698	RLG-----IKLAGPAGDINTGPWATFIRKFHNPRDWDPLVNLFLPLPSAGRQI-						
Qy	8	RLGPPGRMLQORGIEDEBGRVVVVFNPNSRREWAA						
Db	698	RLG-----IKLAGPAGDINTGPWATFIRKFHNPRDWDPLVNLFLPLPSAGRQI-						
Qy	58	SIVTHIQWMAVHTREL--LSSGQMPFSQLPFWVOLWPGKRYRAADSSEWEIADHGQDSM	115					
Db	751	---HLAMASEKFETPELESAVRAAMEAADVPL---FQSALSVMWLENGIVTDM	801					
Qy	116	EQLVLT 121						
Db	802	ANFALS 807						
		RESULT 110						
		US-08-219-62B-12						
		Sequence 12, Application US/08219262B						
		PATENT NO. 5788910						
		GENERAL INFORMATION:						
		APPLICANT: VAKHARIA, VIKRAM						
		APPLICANT: SNYDER, DAVID B						
		APPLICANT: MENGEI-WHERST, STEPHANIE A						
		TITLE OF INVENTION: CHIMERIC INFECTIOUS BURSAL DISEASE VIRUS						
		TITLE OF INVENTION: CDNA CLONES, EXPRESSION PRODUCTS AND VACCINES BASED						
		NUMBER OF SEQUENCES: 15						
		CORRESPONDENCE ADDRESS:						
		ADDRESSEE: OBLON, SPIWAK, MCCLELLAND, MATER & NEUSTADT						
		STREET: 1755 S. JEFFERSON DAVID HIGHWAY, FOURTH FLOOR						
		CITY: ARLINGTON						
		STATE: VIRGINIA						
		COUNTRY: USA						
		ZIP: 22202						
		COMPUTER READABLE FORM:						
		MEDIUM TYPE: Floppy disk						
		COMPUTER: IBM PC compatible						
		OPERATING SYSTEM: PC-DOS/MS-DOS						
		SOFTWARE: Patentin Release #1.0, Version #1.30						
		CURRENT APPLICATION DATA:						
		APPLICATION NUMBER: US/08/219,262B						
		FILING DATE: 29-MAR-1994						
		CLASSIFICATION: 435						
		ATTORNEY/AGENT INFORMATION:						
		NAME: OBLON, NORMAN F						
		REGISTRATION NUMBER: 24,618						
		REFERENCE/DOCKET NUMBER: 2747-047-27						
		TELECOMMUNICATION INFORMATION:						
		TELEPHONE: (703) 413-3000						
		TELEFAX: (703) 413-2220						
		TELEX: 248855 OPAT UR						
		INFORMATION FOR SEQ ID NO: 12:						
		SEQUENCE CHARACTERISTICS:						
		LENGTH: 1012 amino acids						
		TYPE: amino acid						
		TOPOLOGY: linear						
		MOLECULE TYPE: protein						
		US-08-219-262B-12						
Query		Query Match	10,18;	Score 69;	DB 1;	Length 1012;		
		Best Local Similarity	23,8%;	Pred. No. 11;				
		Matches	30;	Conservative	18;	Mismatches	50;	Indels 28; Gaps 5;
Qy	8	RIGVPPGRMLQORGIEDEBGRVVVVFNPNSRREWAA						
Db	698	RLG-----IKLAGPAGDINTGPWATFIRKFHNPRDWDPLVNLFLPLPSAGRQI-						
Qy	58	SIVTHIQWMAVHTREL--LSSGQMPFSQLPFWVOLWPGKRYRAADSSEWEIADHGQDSM	115					
Db	751	---HLAMASEKFETPELESAVRAAMEAADVPL---FQSALSVMWLENGIVTDM	801					
Qy	116	EQLVLT 121						

Db 802 ANFALS 807

RESULT 12  
US-09-031-655-9  
: Sequence 9. Application US/09031655

Patent No. 6017759

11  
-03-1655-1  
Sequence No. 1, Application US/09031655  
ENT No. 6017759  
GENERAL INFORMATION:  
APPLICANT: VAKHARIA, VIKRAM  
APPLICANT: SNIDER, DAVID B  
APPLICANT: MENGEI-WHESAT, STEPHANIE A  
TITLE OF INVENTION: CHIMERIC INFECTIOUS BURSAL DISEASE VIRUS  
TITLE OF INVENTION: CDNA CLONES, EXPRESSION PRODUCTS AND VACCINES BASED  
TITLE OF INVENTION: THEREON  
NUMBER OF SPECIMENS: 15

FIGURE 1  
TITLE OF INVENTION: CHIMERIC INFECTIOUS BURSAL DISEASE VIRUS  
TITLE OF INVENTION: CDNA CLONES, EXPRESSION PRODUCTS AND VACCINES BASED  
TITLE OF INVENTION: THEREON  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAYER & NEUSTADT  
STREET: 1755 S. JEFFERSON DAVID HIGHWAY, FOURTH FLOOR  
CITY: ARLINGTON  
STATE: VIRGINIA

ADDRESS: 1755 S. JEFFERSON DAVID HIGHWAY, FOURTH FLOOR  
CITY: ARLINGTON STATE: VIRGINIA  
COUNTRY: USA  
ZIP: 22202

COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent-in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/031,655

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release 1.0, version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/031,655  
FILING DATE:  
CLASSIFICATION:  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 08/219,262  
FILING DATE: 29-MAR-1994  
ATTORNEY/AGENT INFORMATION:

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/219,262  
FILING DATE: 29-MAR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: OBLON, NORMAN F  
REGISTRATION NUMBER: 24,618  
REFERENCE/DOCKET NUMBER: 2747-047-27  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 413-3000  
TELEFAX: (703) 413-2220  
TELEX: 248855 OBLON UR  
INFORMATION FOR SEO ID NO: 1

MOLECULE TYPE: protein  
; ORIGIN: source;  
; ORGANISM: Infectious bursal disease virus  
; STRAIN: 002-73  
IT-S-09-031-655-9

-----  
ORGANISM: Infectious bursal disease virus  
; STRAIN: GLS  
US-09-031-655-1

Query Match 10.1%; Score 69; DB 3; Length 1012;  
 Best Local Similarity 24.6%; pred. No. 11;  
 Matches 31; Conservative 16; Mismatches 51; Indels 28; Gaps 5  
 Qy 8 RIGWPPGRWQRCIYDEEGRRTWVYVWRFPSRREWAR-----ASQCSRVP 57

Matches	Similarity	Conservative	18;	Mismatches	50:	Indels	28;	Gaps
Best Local	23.8%			pred.	0.11;			
QY	8	RIGUPGRWIMQRPVYEDGETWVWVWVNPBRREMAR-----ASGDSRYPE	57					
Db	698	RLG-----LKLAGPAGFADYNTGPWATFIKRPHNPRMDRDLRBLNLPYLPNAGROY--	750					
QY	58	SITVHLWOMAHTREL-LSSQMPFSOLPAWQOLYPGKRYRADSSFWELADHQISM	115					
Db	751	---HLAMAAEFKPEPELSAYRAMRAASDPL---FOQASLWMFWEENIVDM	801					

RESULT 13  
US-09-031-655-12  
; Sequence 12, Application US/090316555  
; Patent No. 6017759

## GENERAL INFORMATION:

APPLICANT: VAKHARIA, VIKRAM

APPLICANT: SNYDER, DAVID B

APPLICANT: MENGEI-WHERSAT, STEPHANIE A

TITLE OF INVENTION: CHIMERIC INFECTIONS BURSAL DISEASE VIRUS

TITLE OF INVENTION: DNA CLONES, EXPRESSION PRODUCTS AND VACCINES BASED

TITLE OF INVENTION: THEREON

NUMBER OF SEQUENCES: 15

CORRESPONDENCE ADDRESS:

ADDRESSEE: OBLON, SPTYAK, MCCLELLAND, MAIER &amp; NEUSTADT

CITY: ARLINGTON

STATE: VIRGINIA

COUNTRY: USA

ZIP: 22202

COMPUTER READABLE FORM:

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/031,655

FILING DATE:

CLASSIFICATION:

PRIORITY APPLICATION DATA:

APPLICATION NUMBER: US 08/219,262

FILING DATE: 22-MAR-1994

ATTORNEY/AGENT INFORMATION:

NAME: OBLON, NORMAN F

REGISTRATION NUMBER: 24,618

REFERENCE/DOCKET NUMBER: 2747-047-27

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 413-3000

TELEFAX: 248955 OPAT UR

ATTORNEY/AGENT ID NO: 12:

SEQUENCE CHARACTERISTICS:

LENGTH: 1012 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-031-655-12

Query Match 10.1%; Score 69; DB 3; Length 1012;

Best Local Similarity 23.8%; Pred. No. 11; Mismatches 50; Indels 28; Gaps 5;

Matches 30; Conservative 18; Mismatches 50; Indels 28; Gaps 5;

QY 8 RLGVPPGRWLRIOPGTIEDEDEGRWVTVVRENPRREWAR-----ASGSRVEP 57

Db 698 RLG-----LKLPGDAFDVWGPWATFPRKPRHPRDRFLYLPNLPNAGROY-- 750

QY 58 SITVILWQAVHIVREL--LSSGQNPESQPLAWVOLYPSGRKYRADSFVWIAHQDSM 115

Db 751 ---HLAMASKEFKEPELSAVYAMRAASVDP-----FOSALVWMLLENGTIVDM 801

QY 116 EQLVIT 121

Db 802 ANFALS 807

RESULT 14

US-08-708-541A-34

Sequence 34, Application US/08708541A

Patient No. 5871744

GENERAL INFORMATION:

APPLICANT: VAKHARIA, VIKRAM N.

APPLICANT: SNYDER, DAVID B

APPLICANT: MENGEI-WHERSAT, STEPHANIE A

TITLE OF INVENTION: A METHOD FOR GENERATING BIRNAVIRUS FROM

NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESS:

ADDRESSEE: NIKAIKO, MARMELSTEIN, MURRAY &amp; ORAM LLP

STREET: 655 Fifteenth Street, N. W.

STREET: Suite 330 - G Street, Lobby

CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20005-5701

COMPUTER READABLE FORM:

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/708,541A

FILING DATE:

CLASSIFICATION: 4224

ATTORNEY/AGENT INFORMATION:

NAME: KITTS, MONICA C.

REGISTRATION NUMBER: 36,105

REFERENCE/DOCKET NUMBER: P8172-6002

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202/638-5000

TELEFAX: 202/638-4810

INFORMATION FOR SEQ ID NO: 34:

SEQUENCE CHARACTERISTICS:

LENGTH: 1013 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-708-541A-34

Query Match 9.9%; Score 68; DB 2; Length 1013;

Best Local Similarity 23.1%; Pred. No. 14; Mismatches 61; Indels 10; Gaps 4;

Matches 27; Conservative 19; Mismatches 61; Indels 10; Gaps 4;

QY 8 RLGVPPGRWLRIOPGTIEDEDEGRWVTVVRENPRREWAR-----ASGSRVEP 57

Db 699 RLGMRK-----LAGPGADINPQNPWATFPRKPRHPRDRFLYLPNLPPTAGRQLH 753

QY 67 AVHTRTEILSSGOM--PSQLPAWVOLYPSGRKYRADSFVWIAHQDSM 115

Db 754 ALAASKEFKEPELSAVYAMRAASVDP-----FOSALVWMLLENGTIVDM 801

RESULT 15

US-08-219-262B-7

Sequence 7, Application US/08219262B

Patent No. 5788970

GENERAL INFORMATION:

APPLICANT: VAKHARIA, VIKRAM

APPLICANT: SNYDER, DAVID B

APPLICANT: MENGEI-WHERSAT, STEPHANIE A

TITLE OF INVENTION: CHIMERIC INFECTIONS BURSAL DISEASE VIRUS

TITLE OF INVENTION: DNA CLONES, EXPRESSION PRODUCTS AND VACCINES BASED

TITLE OF INVENTION: THEREON

NUMBER OF SEQUENCES: 15

CORRESPONDENCE ADDRESS:

ADDRESSEE: OBLON, SPTYAK, MCCLELLAND, MAIER &amp; NEUSTADT

CITY: ARLINGTON

STATE: VIRGINIA

COUNTRY: USA

ZIP: 22202

COMPUTER READABLE FORM:

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/219,262B

FILING DATE: 29 MAR-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: OBLON, NORMAN F  
 REGISTRATION NUMBER: 24,618  
 REFERENCE/DOCKET NUMBER: 2747-047-27  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (703) 413-3000  
 TELEX: (703) 413-2220  
 TELEX: 24855 OPAT UR  
 INFORMATION FOR SEQ ID NO: 7:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1012 amino acids  
 TYPE: amino acid  
 STRANDEDNESS:  
 TOPOLOGY: unknown  
 MOLECULE TYPE: protein  
 ORIGINAL SOURCE:  
 ORGANISM: Infectious bursal disease virus  
 STRAIN: 52/70  
 US-08-219-262B-7

Qy	Ay	Match	Score	DB	Length	DB1	Length1
Qy	8	RIGVPPGRWLRQIRGIVDEERGRTWTVVRENPSREWARASQGSRVEPSITVHLWOMA	67				
Db	698	RIKG---LKLPGPARDGIVNGPWWATIKRKPRAFPNDR-----	733				
Qy	68	VHTRELLSSGOMPFSOLPAWQLYP--GRKRAADSSFWIADHGQIDS	114				
Db	734	-----LPYLNLP--YLPNNAGRQYHLLAMAAS-BFKDPELES	767				

Search completed: November 29, 2001, 04:03:31  
Job time: 333 sec

Job time: 333 sec

Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 29, 2001, 03:59:59 ; Search time 18:53 Seconds

title: US-09-526-329-39

perfect score: 685

sequence: 1 MASEASVRLGVPPGRMIIQR.....HQQIDSMEQLVLYQPERKD 128

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

PIR\_65:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	176.5	25.8	S78532	MTCP-1 protein splice form Bl - human
2	149	21.8	S138286	T cell leukemia/ly
3	79.5	11.6	S11839	vird2 protein - Ag
4	79.5	11.6	S37763	vird2 protein - Ag
5	79.5	11.6	S03020	dopamine beta-mono
6	79	11.5	C83240	probable transcript
7	75.5	11.0	E83796	two-component sens
8	74.5	10.9	S436	vird2 protein Ag
9	73.5	10.7	J00137	hypothetical
10	73	10.7	G83175	30.5
11	72.5	10.6	B25063	probable metallo- $\alpha$
12	72.5	10.6	T34689	hypothetical vird2
13	72.5	10.6	S138286	probable iron-sulf
14	70	10.2	S367	probable mbtb prot
15	70	10.2	C82391	N-ethylmaleimide r
16	70	10.2	T07824	RESULT
17	69.5	10.1	S34370	2
18	69.5	10.1	T41647	138286
19	69.5	10.1	363	T cell leukemia/lymphoma protein T-cell - human
20	69.5	10.1	433	N; Alternate names: T-cell leukemia-related protein T-cell
21	69	10.1	1927	C; Species: Homo sapiens (man)
22	69	10.1	363	C; Date: 09-Mar-1996 #sequence_revision 09-Mar-1996
23	69	10.1	1012	C; Accession: I-18286; S1138
24	68	10.1	1747	R; Virgilio, L.; Narducci, M.G.; Itobe, M.; Billis, L.G.; Cooper, M.D.; Croce, C
25	67.5	9.9	1013	Proc. Natl. Acad. Sci. U.S.A. 91, 12530-12534, 1994
26	67.5	9.9	762	A; Title: Identification of the T-cell gene involved in T-cell malignancies.
27	67.5	9.9	1012	A; Reference number: I-18286; MUID:95107991
28	67	9.8	1148	A; Accession: I-18286; MUID:95107991
29	66.5	9.7	2	A; Molecule type: mRNA
				A; Residues: 1-114 <RES>
				A; Cross-references: PDB:1-18286; MUID:95107991; PDB:1-18286; MUID:95107991
				30 66.5 9.7 1209 2 H05839
				31 66 9.7 237 2 C8183
				32 66 9.6 363 2 T3072
				33 66 9.6 536 2 D83622
				34 66 9.6 993 1 GNN98
				35 66 9.6 1001 2 C86181
				36 66 9.6 1012 1 GNN9R
				37 66 9.6 1012 1 GNN9C
				38 66 9.6 1170 2 H71295
				39 66 9.6 1175 2 T25634
				40 65.5 9.6 327 2 T15594
				41 65.5 9.6 350 2 D83102
				42 65.5 9.6 764 2 D84847
				43 65 9.5 286 2 SA48201
				44 65 9.5 352 2 S76078
				45 65 9.5 406 2 C83867

ALIGNMENTS



Db 377 TDKCTQALAPPSGHIIFASOLHHLTGRKVTVLVR--DGREWELVNQDNHYSPP---- 428  
 Qy 63 IWMQAVHRELLSSGOMPFSQIARWMLYPCR----KYRADSSFWELADHQDQSME 116  
 Db 429 ------HQEI-----RMLKVVSVHPGDVILITSCYNTEDR--ELATVGGFILE 471  
 Qy 117 QLVHTY 122  
 Db 472 EMCVNY 477

RESULT 6  
 C03240 probable transcription regulator PA3249 [imported] - *Pseudomonas aeruginosa* (strain PA01  
 C;Species: *Pseudomonas aeruginosa*  
 C;Accession: C03240  
 R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.; Bir  
 adam, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,  
 Lory, S.; Olson, M.V.  
 Nature 405, 559-564, 2000  
 A;Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic patho  
 A;Reference number: A82950; MUID:2043737  
 A;Accession: C03240  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-238 <STO>  
 A;Experimental source: strain PA01  
 C;Genetics:  
 A;Gene: PA3249

RESULT 7  
 E883796 Query Match 11.5%; Score 79; DB 2; Length 238;  
 Best Local Similarity 23.9%; Pred. No. 0.75; Pred. 25.6%;  
 Matches 28; Conservative 17; Mismatches 46; Indels 26; Gaps 3;  
 C;Species: *Bacillus halodurans*  
 C;Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 31-Dec-2000  
 R;Accession: E883796  
 R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hirai  
 Nucleic Acids Res. 28, 4317-4321, 2000  
 A;Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and  
 A;Reference number: A83650; MUID:20263314  
 A;Accession: E883796  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-476 <STO>  
 A;Cross-references: GB:AP001511; GB:BA000004; NID:910173727; PIDN:BAR04992.1; GSPDB:GN001  
 C;Genetics:  
 A;Gene: BH1173

RESULT 8  
 S06884 Query Match 10.9%; Score 74.5; DB 2; Length 436;  
 Best Local Similarity 25.6%; Pred. 4.7%;  
 Matches 34; Conservative 17; Mismatches 43; Indels 39; Gaps 8;  
 C;Species: *Agrobacterium rhizogenes*  
 C;Accession: S06884  
 R;Hirayama, T.; Muranaka, T.; Ohawa, H.; Oka, A.  
 Mol. Gen. Genet. 213, 229-237, 1988  
 A;Title: Organization and characterization of the virCD genes from *Agrobacterium rhiz*  
 A;Reference number: S06881; MUID:89039712  
 A;Accession: S06884  
 A;Molecule type: DNA  
 A;Residues: 1-436 <HIR>  
 A;Cross-references: EMBL:X12867; NID:938995; PIDN:CAR31351.1; PID:939000  
 A;Gene: virD2  
 C;Genetics:  
 A;Genome: plasmid

RESULT 9  
 J00137 Query Match 10.7%; Score 73.5; DB 2; Length 261;  
 Best Local Similarity 27.0%; Pred. No. 3.2%;  
 Matches 30; Conservative 11; Mismatches 41; Indels 29; Gaps 6;  
 C;Species: *Pseudomonas aeruginosa*  
 C;Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 07-Jun-1999  
 R;Accession: J00137  
 R;Kato, J.; Chu, L.; Kitano, K.; Devault, J.D.; Kimbara, K.; Chakrabarty, A.M.; Misra  
 Gene 84, 31-38, 1989  
 A;Title: Nucleotide sequence of a regulatory region controlling alginate synthesis in  
 A;Reference number: J00132; MUID:90108714  
 A;Accession: J00137  
 A;Status: translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-261 <KAT>  
 A;Note: 3-Met could also be the initiator  
 C;Genetics:  
 A;Start codon: GTG

Query Match 11.0%; Score 75.5; DB 2; Length 476;  
 Best Local Similarity 27.7%; Pred. No. 4.1%;  
 Matches 36; Conservative 16; Mismatches 53; Indels 25; Gaps 7;

Qy 19 QRPGLYDEEGKTVWVVFNRNSRRE-WARASGSRVPEPSTIVLWQMVHRELLS 76  
 Db 115 QSRGRYELSYNQATIVWVTKFNSFGREGYLISYMDTYDRAVNRWLRILYIL-LLSS 173

Query Match 10.7%; Score 73.5; DB 2; Length 261;  
 Best Local Similarity 27.0%; Pred. No. 3.2%;  
 Matches 30; Conservative 11; Mismatches 41; Indels 29; Gaps 6;  
 C;Species: *Pseudomonas aeruginosa*  
 C;Accession: J00137  
 R;Kato, J.; Chu, L.; Kitano, K.; Devault, J.D.; Kimbara, K.; Chakrabarty, A.M.; Misra  
 Gene 84, 31-38, 1989  
 A;Title: Nucleotide sequence of a regulatory region controlling alginate synthesis in  
 A;Reference number: J00132; MUID:90108714  
 A;Accession: J00137  
 A;Status: translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-261 <KAT>  
 A;Note: 3-Met could also be the initiator  
 C;Genetics:  
 A;Start codon: GTG



Query Match 10.6%; Score 72.5; DB 2; Length 1414;  
 Best Local Similarity 26.1%; Pred. No. 32; Mismatches 9; Indels 21; Gaps 2;  
 Matches 30; Conservative

QY 5 ASVRIGVPP-----PGRWIKQIOPGIVYDEDEGRFWVTVVRFNPSSR 44  
 Db 862 ASVPYGVPPNPACRWVADSGCDPDPWVAGELWVSGRGITARGYGRGRELTAERWEHGR 921  
 QY 45 EMARASOGQRYEFSIVHLWOMAHITRELLSGQMPFSQLPAWQLYPGKRYRA 99  
 Db 922 TWYRIGDLARYWHDGTLEFVGRADH-RVKISQYRVELGETEAMQLRPGVHAA 975

RESULT 14

C82391 N-ethylmaleimide reductase VCA0993 [imported] - *Vibrio cholerae* (strain N16961 serogroup C;Species: *Vibrio cholerae* C;Genetics: C;Accession: C82391 C;Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001

R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Givn, M.L.; Dodson, R.J.; chardson, D.; Brimley, M.D.; Vaishnav, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, F.; L.; R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M. Nature 406, 477-483, 2000

A;Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*. A;Reference number: AB2035; MUID:2046833

A;Accession: C82391 A;Status: preliminary A;Molecule type: DNA A;Residues: 1-367 <HET>

A;Cross-references: GB:AB04426; GB:E003853; NID:965831; PIDN:AM96889.1; GSPDB:GN001 A;Experimental source: serogroup 01; strain N16961; biotype El Tor

A;Gene: VCA0993 A;Map position: 2

C;Superfamily: NADPH dehydrogenase chain OYE2

Query Match 10.2%; Score 70; DB 2; Length 367;  
 Best Local Similarity 24.8%; Pred. No. 11; Mismatches 26; Conservative 18; Mismatches 33; Indels 28; Gaps 7;  
 Matches 26; Conservative

QY 21 RGIYEDERGRFTWVWRFNPSSREWARSOGQRYEPSPITVHLWOMAHITRELLSGQMP 80  
 Db 71 FGVTYDDEOTSGKVTI-----QAVKQOGA---AMFCQIWHVGRYSHPWFQKGQLP 117  
 QY 81 FSQLPAWQVQLPGRKRAADSSWEIAD--HGO-IDSMEQIVIT 121  
 Db 118 RA-PSA-----LKPVETQW-IADEGNGQMDVPEPAMT 151

RESULT 15

T07824 1.4-alpha-glucan branching enzyme (EC 2.4.1.18) I (clone sbel7) - potato (fragment) C;Species: *Solanum tuberosum* (potato) C;Accession: 14-May-1999 #sequence\_revision 14-May-1999 #text\_change 21-Jul-2000 C;Date: T07824 R;Khoshnoodi, J.; Blenow, A.; Ek, B.; Rask, L.; Larsson, H. Eur. J. Biochem. 242, 148-155, 1996

A;Title: The multiple forms of starch branching enzyme I in *Solanum tuberosum*. A;Reference number: Z16155; MUID:97112484

A;Accession: T07824 A;Status: preliminary; translated from GB/EMBL/DBJ A;Molecule type: mRNA A;Residues: 1-830 <HET>

A;Cross-references: EMBL:Y08786; NID:91621011; PIDN:CAA70038.1; PID:91621012 A;Experimental source: cv. Dianella; cell line Dianella C;Genetics: A;Gene: sbel7 C;Function: C;Description: catalyzes the transglycosylation of a terminal segment of a 1,4-alpha-D-g C;Superfamily: 1,4-alpha-glucan branching enzyme

C;Keywords: glycogen/starch biosynthesis; glycosyltransferase; hexosyltransferase

Query Match 10.2%; Score 70; DB 2; Length 830;  
 Best Local Similarity 24.2%; Pred. No. 30; Mismatches 31; Conservative 17; Mismatches 34; Indels 46; Gaps 8;  
 Matches 31; Conservative

QY 17 WIQRPGIVYDEDEGRFWVTVVRFNPSSREWARSOGQRYEPSPITVHLWOMAHITRELLSGQMP 80  
 Db 545 WIDFP-----REGNNW---SYDKRCRQWNLADSHLRK--FMAFDRAMNSLDEKF 591  
 QY 73 -LISQCMPEF-----OLPAVWOLYPSR-----KYRAA-DSSF 103  
 Db 592 SFLASGKQIVSSMDDDKWVVFGRDLVVFVFNPHKNTVEGYKVGCQLPGKRYVALDSDA 651

QY 104 WELADHGQ 111  
 Db 652 WEGGHGR 659

Search completed: November 29, 2001, 04:04:10  
 Job time: 231 sec

Fri Nov 30 11:03:11 2001

us-09-526-329-39.rpr

Page 6

GenCore version 4.5									
Copyright (c) 1993 - 2000 Compugen Ltd.									
<b>OM protein - protein search, using sw model</b>									
Run on: November 29, 2001, 04:03:14 ; Search time 9.89 seconds									
Title: US-09-526-329-39									
Perfect score: 695									
Sequence: 1 MASEASVRLGVPPGRWIOR. .... HQQIDSMEDQLVLYQPERKD 128									
Scoring table: BLOSUM62									
GapOp 10.0 , GapExt 0.5									
Searched: 100059 seqs, 36664827 residues									
Total number of hits satisfying chosen parameters: 100059									
Minimum DB seq length: 0									
Maximum DB seq length: 200000000									
Post-processing: Minimum Match 0%									
Maximum Match 100%									
Listing first 45 summaries									
Database : SwissProt_39:*									
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.									
SUMMARIES									
Result No.	Score	Query Match	Length	DB ID	Description	RESULT	1	ALIGNMENTS	
1	685	100.0	128	1 TCLB_HUMAN	095988 homo sapien	TCLB_HUMAN			
2	178.5	26.1	116	1 TBLB_MOUSE	P56940 mus musculus	TCLB_HUMAN	STANDARD;		
3	178.5	26.1	120	1 TBLB_MOUSE	P56844 mus musculus	TCLB_HUMAN	PRT;	128 AA.	
4	178	26.0	107	1 MTC2_HUMAN	060945 mus musculus	095988; 30-MAY-2000 (rel. 39, Created)	DT		
5	16.5	25.8	175.5	1 TBLB_MOUSE	P56278 homo sapien	30-MAY-2000 (rel. 39, Last sequence update)	DT		
6	15.5	25.6	122	1 TBLB_MOUSE	P56280 mus musculus	T-CELL LEUKEMIA/LYMPHOMA PROTEIN 1B (TCLB ONCOGENE)	DE		
7	15.9	24.7	116	1 TBLB_MOUSE	P56280 mus musculus	(SYNCTIOTROPHOBLAST-SPECIFIC PROTEIN) (SYN-1)	DE		
8	16.35	23.9	121	1 TBLB_MOUSE	P56285 mus musculus	TCLB_HUMAN	GN		
9	15.8.5	23.1	117	1 TBLB_MOUSE	P56841 mus musculus	Homo sapiens (Human)	OS		
10	14.9	21.8	114	1 TBLA_HUMAN	P56279 homo sapien	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	OC		
11	79.5	11.6	447	1 VTD2_AGR75	P18592 agrobacteri	NCBI_TaxID=9606; NCBI_TaxID=9606;	OX		
12	79.5	11.6	603	1 DOPD_HUMAN	P09172 homo sapien	SEQUENCE FROM N.A.	RN		
13	74.5	10.9	436	1 VTD2_AGR75	P13462 agrobacteri	SEQUENCE FROM N.A.	RN		
14	72.5	10.6	424	1 VTD2_AGR76	P06668 agrobacteri	"A syncytiotrophoblast-specific gene Syn-1 cloned from human syncytiotrophoblast subtracted cDNA library.";	RT		
15	70	10.2	861	1 GLB2B_SOLTU	P03924 solanum tub	Submitted (MAR-1999) to the EMBL/Genbank/DDBJ databases.	RT		
16	69.5	10.1	366	1 VIL1_HSV6U	P01349 human herpesvirus 1	-- TISSUE SPECIFICITY: EXPRESSED IN A VARIETY OF TISSUES.	CC		
17	69.5	10.1	775	1 LYSA4_EMENTI	Q24121 emericella	REARRANGEMENTS OF THE 14Q32.1 REGION.	CC		
18	69.5	10.1	1411	1 Y397_HUMAN	O15040 homo sapien	-- DISEASE: ACTIVATED IN CHRONIC T-CELL LEUKEMIAS (T-CELL) CARRYING	CC		
19	69	10.1	1012	1 P0132_HDVA	P03864 avian infect	-- SIMILARITY: BELONGS TO THE TCELL FAMILY.	CC		
20	67.5	9.9	762	1 P0M74 YEAST	P06971 saccharomyces cerevisiae	This SWISS-PROT entry is copyright. It is produced through a collaborative effort between the Swiss Institute of Bioinformatics and the EMBL outstation of the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by for commercial entities requires a license agreement. (See <a href="http://www.isb-sib.ch/announce.html">http://www.isb-sib.ch/announce.html</a> or send an email to license@isb-sib.ch).	CC		
21	67.5	9.9	1012	1 P0L5_TBDV	P25220 avian infect	EMBL; AF110466; ADI16997.1; --	CC		
22	67	9.8	125	1 D8SC_COXBD	P51055 coxiella bu	EMBL; AF110465; ADI16996.1; --	CC		
23	67	9.8	3057	1 CALC_MOUSE	P15480 avian infect	HSSP; P5628; 1A1X.	CC		
24	66.5	9.7	595	1 YG28_BPML5	P060847 mus musculus	MM; 603769; --	CC		
25	66.5	9.7	1490	1 CRK7_HUMAN	Q52355 mycobacteri	InterPro; IPR002709; P015575; TBL1_MTC2P1; 1.	CC		
26	66	9.6	326	1 TPR4_BACER	Q9Y4V4 homo sapien	PROTO-ONCOGENE: Chromosomal translocation.	CC		
27	66	9.6	535	1 ARS_PSEAE	P37247 bacteroides pseudos	SEQUENCE 128 AA; 14446 MW; 6B2CF407884D CRC64;	CC		
28	66	9.6	993	1 P0LS_IBDVC	P25220 avian infect	DR			
29	66	9.6	1012	1 P0LS_IBDVS	P15480 avian infect	DR			
30	66	9.6	1012	1 P0LS_IBDVS	P2351 avian infect	DR			
31	66	9.6	1170	1 DPP3A_TREPA	Q83675 treponema p	DR			
32	65.5	9.6	808	1 SVB_SYNPI	P74764 synchococc	DR			
33	65.5	9.6	877	1 SIA_THIFP	P056273 thiobacillus	proto-oncogene: Chromosomal translocation.	CC		



-1 - SIMILARITY: BELONGS TO THE TCL1 FAMILY.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC

CC EMBL; U32332; AAC52444; 1; -.

CC DR; HSSP; P56278; 1AIX.

CC DR; MGD; MGI-10269; Mtcp1.

CC DR; InterPro; IPR002709; TCL1\_MTCP1.

CC DR; Pfam; PF01840; TCL1\_MTCP1; 2.

CC DR; Prodom; PD01575; TCL1\_MTCP1; 1.

CC KW Alternative splicing; SEQUENCE 107 AA; 12645 MW; 7996813A182C2308 CRC64; SQ

Query Match 26.0%; Score 178; DB 1; Length 107;

Best Local Similarity 32.5%; Pred. No. 3.6e-12;

Matches 39; Conservative 19; Mismatches 40; Indels 22; Gaps 3;

DR 9 LGVPGRLGMIQPGIYDEGRGRTWVW--VRFNPNSREWWARASGSRSPSTVHLWQ 65

DR 6 VGAPPDHLMWQEGDVEORTWVAVWEESIFKARVQVOVPLGDRPS----- 58

Qy 66 MAVITRELLSSQMPFSQPLWVQKPKRAASSSFHEDLPHQDSDMQLVWYQPE 125

Db 59 -----HLT-----SQLPLMMWOLYPEERYMDNNNSRLWQIQLHHMLVRYQVQELILKLPD 106

DR 59 -----HLT-----SQLPLMMWOLYPEERYMDNNNSRLWQIQLHHMLVRYQVQELILKLPD 106

Qy 15 JUL-1998 (Rel. 36, Created)

DR 15-JUL-1998 (Rel. 36, Last sequence update)

DT 20-AUG-2001 (Rel. 40, Last annotation update)

DE P13 MTCP-1 PROTEIN (MATURE T-CELL PROLIFERATION-1 TYPE B1) (MTCP-1 TYPE B1) (P3JWMP1).

DE MTCBP1 OR C6\_1B.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea.

OX NCBI\_TAXID=9606; [1]

RN 11 SEQUENCE FROM N.A.

RP TISSUE=T-cell;

RC MEDLINE-93308950; PubMed=8361760;

RA Stern M.-H.; Soulier J.; Rosenzweig M.; Nakahara K.; Canki-Klain N.,

RA Aurias A.; Sigaux F.; Kirsch I.R.;

RT "MTCP-1: a novel gene on the human chromosome Xq28 translocated to the T cell receptor alpha/delta locus in mature T cell

RT proliferations"; Oncogene 8:2475-2483(1993).

RN [2] X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).

RX MEDLINE-98188217; PubMed=9520380;

RA Fu Z.-O.; du Bois G.C.; Song S.P.; Kulikovskaya I.; Virgilio L.; Rotstein J.L.; Croce C.M.; Weber I.T.; Harrison R.W.;

RT "Crystal structure of MTCP-1: implications for role of TCL-1 and MTCP-1 in T cell malignancies"; Proc. Natl. Acad. Sci. U.S.A. 95:3413-3418(1998).

CC -1 ALTERNATIVE PRODUCTS: ALTERNATIVE SPlicing produces two short, yet COMPLETELY DIFFERENT PROTEINS. THE LONGER PRODUCT, TYPE-B1, IS KNOWN AS P13 MTCP-1 (AC P56278). THE SHORTER PRODUCT, TYPE-A, IS KNOWN AS P8 MTCP-1 (AC P56277).

CC -1 TISSUE SPECIFICITY: NOT FOUND AT A SIGNIFICANT LEVEL IN ANY TISSUE.

CC -1 DISEASE: DETECTED IN T-CELL LEUKEMIA BEARING A T(X;14) TRANSLOCATION. PLAYS A KEY ROLE IN T-CELL POLYMPHOCYTIC LEUKEMIA.

CC

Query Match 25.8%; Score 176.5; DB 1; Length 107;

Best Local Similarity 33.6%; Pred. No. 5.2e-12;

Matches 43; Conservative 17; Mismatches 43; Indels 25; Gaps 4;

DR 1 MASASVRLGPGRLGMIQPGIYDEGRGRTWVW--VRFNPNSREWWARASGSRSP 57

DR 1 MASASVRLGPGRLGMIQPGIYDEGRGRTWVW--VRFNPNSREWWARASGSRSP 57

Qy 58 SIVVHLWQMAVHTRELLSSQMPFSQPLWVQKPKRAASSSFHEDLPHQDSDMQLVWYQPE 117

Db 58 S-----HLT-----SQLPLMMWOLYPEERYMDNNNSRLWQIQLHHMLVRYQVQELILKLPD 106

Qy 118 LVVYQPE 125

Db 99 LLLKLKLPD 106

Query Match 25.8%; Score 176.5; DB 1; Length 107;

Best Local Similarity 33.6%; Pred. No. 5.2e-12;

Matches 43; Conservative 17; Mismatches 43; Indels 25; Gaps 4;

DR 1 MASASVRLGPGRLGMIQPGIYDEGRGRTWVW--VRFNPNSREWWARASGSRSP 57

DR 1 MASASVRLGPGRLGMIQPGIYDEGRGRTWVW--VRFNPNSREWWARASGSRSP 57

Qy 59 SIVVHLWQMAVHTRELLSSQMPFSQPLWVQKPKRAASSSFHEDLPHQDSDMQLVWYQPE 117

Db 58 S-----HLT-----SQLPLMMWOLYPEERYMDNNNSRLWQIQLHHMLVRYQVQELILKLPD 106

Qy 118 LVVYQPE 125

Db 99 LLLKLKLPD 106

Query Match 25.8%; Score 176.5; DB 1; Length 107;

Best Local Similarity 33.6%; Pred. No. 5.2e-12;

Matches 43; Conservative 17; Mismatches 43; Indels 25; Gaps 4;

DR 1 MASASVRLGPGRLGMIQPGIYDEGRGRTWVW--VRFNPNSREWWARASGSRSP 57

DR 1 MASASVRLGPGRLGMIQPGIYDEGRGRTWVW--VRFNPNSREWWARASGSRSP 57

Qy 58 SIVVHLWQMAVHTRELLSSQMPFSQPLWVQKPKRAASSSFHEDLPHQDSDMQLVWYQPE 117

Db 58 S-----HLT-----SQLPLMMWOLYPEERYMDNNNSRLWQIQLHHMLVRYQVQELILKLPD 106

Qy 118 LVVYQPE 125

Db 99 LLLKLKLPD 106

Query Match 25.8%; Score 176.5; DB 1; Length 107;

Best Local Similarity 33.6%; Pred. No. 5.2e-12;

Matches 43; Conservative 17; Mismatches 43; Indels 25; Gaps 4;

DR 1 MASASVRLGPGRLGMIQPGIYDEGRGRTWVW--VRFNPNSREWWARASGSRSP 57

DR 1 MASASVRLGPGRLGMIQPGIYDEGRGRTWVW--VRFNPNSREWWARASGSRSP 57

Qy 59 SIVVHLWQMAVHTRELLSSQMPFSQPLWVQKPKRAASSSFHEDLPHQDSDMQLVWYQPE 117

Db 58 S-----HLT-----SQLPLMMWOLYPEERYMDNNNSRLWQIQLHHMLVRYQVQELILKLPD 106

Qy 118 LVVYQPE 125

Db 99 LLLKLKLPD 106

Query Match 25.8%; Score 176.5; DB 1; Length 107;

Best Local Similarity 33.6%; Pred. No. 5.2e-12;

Matches 43; Conservative 17; Mismatches 43; Indels 25; Gaps 4;

DR 1 MASASVRLGPGRLGMIQPGIYDEGRGRTWVW--VRFNPNSREWWARASGSRSP 57

DR 1 MASASVRLGPGRLGMIQPGIYDEGRGRTWVW--VRFNPNSREWWARASGSRSP 57

Qy 59 SIVVHLWQMAVHTRELLSSQMPFSQPLWVQKPKRAASSSFHEDLPHQDSDMQLVWYQPE 117

Db 58 S-----HLT-----SQLPLMMWOLYPEERYMDNNNSRLWQIQLHHMLVRYQVQELILKLPD 106

Qy 118 LVVYQPE 125

Db 99 LLLKLKLPD 106

Query Match 25.8%; Score 176.5; DB 1; Length 107;

Best Local Similarity 33.6%; Pred. No. 5.2e-12;

Matches 43; Conservative 17; Mismatches 43; Indels 25; Gaps 4;

DR 1 MASASVRLGPGRLGMIQPGIYDEGRGRTWVW--VRFNPNSREWWARASGSRSP 57

DR 1 MASASVRLGPGRLGMIQPGIYDEGRGRTWVW--VRFNPNSREWWARASGSRSP 57

Qy 59 SIVVHLWQMAVHTRELLSSQMPFSQPLWVQKPKRAASSSFHEDLPHQDSDMQLVWYQPE 117

Db 58 S-----HLT-----SQLPLMMWOLYPEERYMDNNNSRLWQIQLHHMLVRYQVQELILKLPD 106

Qy 118 LVVYQPE 125

Db 99 LLLKLKLPD 106

Query Match 25.8%; Score 176.5; DB 1; Length 107;

Best Local Similarity 33.6%; Pred. No. 5.2e-12;

Matches 43; Conservative 17; Mismatches 43; Indels 25; Gaps 4;

DR 1 MASASVRLGPGRLGMIQPGIYDEGRGRTWVW--VRFNPNSREWWARASGSRSP 57

DR 1 MASASVRLGPGRLGMIQPGIYDEGRGRTWVW--VRFNPNSREWWARASGSRSP 57

Qy 59 SIVVHLWQMAVHTRELLSSQMPFSQPLWVQKPKRAASSSFHEDLPHQDSDMQLVWYQPE 117

Db 58 S-----HLT-----SQLPLMMWOLYPEERYMDNNNSRLWQIQLHHMLVRYQVQELILKLPD 106

Qy 118 LVVYQPE 125

Db 99 LLLKLKLPD 106

Query Match 25.8%; Score 176.5; DB 1; Length 107;

Best Local Similarity 33.6%; Pred. No. 5.2e-12;

Matches 43; Conservative 17; Mismatches 43; Indels 25; Gaps 4;

DR 1 MASASVRLGPGRLGMIQPGIYDEGRGRTWVW--VRFNPNSREWWARASGSRSP 57

DR 1 MASASVRLGPGRLGMIQPGIYDEGRGRTWVW--VRFNPNSREWWARASGSRSP 57

Qy 59 SIVVHLWQMAVHTRELLSSQMPFSQPLWVQKPKRAASSSFHEDLPHQDSDMQLVWYQPE 117

Db 58 S-----HLT-----SQLPLMMWOLYPEERYMDNNNSRLWQIQLHHMLVRYQVQELILKLPD 106

Qy 118 LVVYQPE 125

Db 99 LLLKLKLPD 106

Query Match 25.8%; Score 176.5; DB 1; Length 107;

Best Local Similarity 33.6%; Pred. No. 5.2e-12;

Matches 43; Conservative 17; Mismatches 43; Indels 25; Gaps 4;

DR 1 MASASVRLGPGRLGMIQPGIYDEGRGRTWVW--VRFNPNSREWWARASGSRSP 57

DR 1 MASASVRLGPGRLGMIQPGIYDEGRGRTWVW--VRFNPNSREWWARASGSRSP 57

Qy 59 SIVVHLWQMAVHTRELLSSQMPFSQPLWVQKPKRAASSSFHEDLPHQDSDMQLVWYQPE 117

Db 58 S-----HLT-----SQLPLMMWOLYPEERYMDNNNSRLWQIQLHHMLVRYQVQELILKLPD 106

Qy 118 LVVYQPE 125

Db 99 LLLKLKLPD 106

Query Match 25.8%; Score 176.5; DB 1; Length 107;

Best Local Similarity 33.6%; Pred. No. 5.2e-12;

Matches 43; Conservative 17; Mismatches 43; Indels 25; Gaps 4;

DR 1 MASASVRLGPGRLGMIQPGIYDEGRGRTWVW--VRFNPNSREWWARASGSRSP 57

DR 1 MASASVRLGPGRLGMIQPGIYDEGRGRTWVW--VRFNPNSREWWARASGSRSP 57

Qy 59 SIVVHLWQMAVHTRELLSSQMPFSQPLWVQKPKRAASSSFHEDLPHQDSDMQLVWYQPE 117

Db 58 S-----HLT-----SQLPLMMWOLYPEERYMDNNNSRLWQIQLHHMLVRYQVQELILKLPD 106

Qy 118 LVVYQPE 125

Db 99 LLLKLKLPD 106

Query Match 25.8%; Score 176.5; DB 1; Length 107;

Best Local Similarity 33.6%; Pred. No. 5.2e-12;

Matches 43; Conservative 17; Mismatches 43; Indels 25; Gaps 4;

DR 1 MASASVRLGPGRLGMIQPGIYDEGRGRTWVW--VRFNPNSREWWARASGSRSP 57

DR 1 MASASVRLGPGRLGMIQPGIYDEGRGRTWVW--VRFNPNSREWWARASGSRSP 57

Qy 59 SIVVHLWQMAVHTRELLSSQMPFSQPLWVQKPKRAASSSFHEDLPHQDSDMQLVWYQPE 117

Db 58 S-----HLT-----SQLPLMMWOLYPEERYMDNNNSRLWQIQLHHMLVRYQVQELILKLPD 106

Qy 118 LVVYQPE 125

Db 99 LLLKLKLPD 106

Query Match 25.8%; Score 176.5; DB 1; Length 107;

Best Local Similarity 33.6%; Pred. No. 5.2e-12;

Matches 43; Conservative 17; Mismatches 43; Indels 25; Gaps 4;

DR 1 MASASVRLGPGRLGMIQPGIYDEGRGRTWVW--VRFNPNSREWWARASGSRSP 57

DR 1 MASASVRLGPGRLGMIQPGIYDEGRGRTWVW--VRFNPNSREWWARASGSRSP 57

Qy 59 SIVVHLWQMAVHTRELLSSQMPFSQPLWVQKPKRAASSSFHEDLPHQDSDMQLVWYQPE 117

Db 58 S-----HLT-----SQLPLMMWOLYPEERYMDNNNSRLWQIQLHHMLVRYQVQELILKLPD 106

Qy 118 LVVYQPE 125

Db 99 LLLKLKLPD 106

Query Match 25.8%; Score 176.5; DB 1; Length 107;

Best Local Similarity 33.6%; Pred. No. 5.2e-12;

Matches 43; Conservative 17; Mismatches 43; Indels 25; Gaps 4;

DR 1 MASASVRLGPGRLGMIQPGIYDEGRGRTWVW--VRFNPNSREWWARASGSRSP 57

DR 1 MASASVRLGPGRLGMIQPGIYDEGRGRTWVW--VRFNPNSREWWARASGSRSP 57

Qy 59 SIVVHLWQMAVHTRELLSSQMPFSQPLWVQKPKRAASSSFHEDLPHQDSDMQLVWYQPE 117

Db 58 S-----HLT-----SQLPLMMWOLYPEERYMDNNNSRLWQIQLHHMLVRYQVQELILKLPD 106

Qy 118 LVVYQPE 125

Db 99 LLLKLKLPD 106

Query Match 25.8%; Score 176.5; DB 1; Length 107;

Best Local Similarity 33.6%; Pred. No. 5.2e-12;

Matches 43; Conservative 17; Mismatches 43; Indels 25; Gaps 4;

DR 1 MASASVRLGPGRLGMIQPGIYDEGRGRTWVW--VRFNPNSREWWARASGSRSP 57

DR 1 MASASVRLGPGRLGMIQPGIYDEGRGRTWVW--VRFNPNSREWWARASGSRSP 57

Qy 59 SIVVHLWQMAVHTRELLSSQMPFSQPLWVQKPKRAASSSFHEDLPHQDSDMQLVWYQPE 117

Db 58 S-----HLT-----SQLPLMMWOLYPEERYMDNNNSRLWQIQLHHMLVRYQVQELILKLPD 106

Qy 118 LVVYQPE 125

Db 99 LLLKLKLPD 106

Query Match 25.8%; Score 176.5; DB 1; Length 107;

Best Local Similarity 33.6%; Pred. No. 5.2e-12;

Matches 43; Conservative 17; Mismatches 43; Indels 25; Gaps 4;

DR 1 MASASVRLGPGRLGMIQPGIYDEGRGRTWVW--VRFNPNSREWWARASGSRSP 57

DR 1 MASASVRLGPGRLGMIQPGIYDEGRGRTWVW--VRFNPNSREWWARASGSRSP 57

Qy 59 SIVVHLWQMAVHTRELLSSQMPFSQPLWVQKPKRAASSSFHEDLPHQDSDMQLVWYQPE 117

Db 58 S-----HLT-----SQLPLMMWOLYPEERYMDNNNSRLWQIQLHHMLVRYQVQELILKLPD 106

Qy 118 LVVYQPE 125

Db 99 LLLKLKLPD 106

Query Match 25.8%; Score 176.5; DB 1; Length 107;

Best Local Similarity 33.6%; Pred. No. 5.2e-12;

Matches 43; Conservative 17; Mismatches 43; Indels 25; Gaps 4;

DR 1 MASASVRLGPGRLGMIQPGIYDEGRGRTWVW--VRFNPNSREWWARASGSRSP 57

DR 1 MASASVRLGPGRLGMIQPGIYDEGRGRTWVW--VRFNPNSREWWARASGSRSP 57

Qy 59 SIVVHLWQMAVHTRELLSSQMPFSQPLWVQKPKRAASSSFHEDLPHQDSDMQLVWYQPE 117

Db 58 S-----HLT-----SQLPLMMWOLYPEERYMDNNNSRLWQIQLHHMLVRYQVQELILKLPD 106

Qy 118 LVVYQPE 125

Db 99 LLLKLKLPD 106

Query Match 25.8%; Score 176.5; DB 1; Length 107;

Best Local Similarity 33.6%; Pred. No. 5.2e-12;

Matches 43; Conservative 17; Mismatches 43; Indels 25; Gaps 4;

DR 1 MASASVRLGPGRLGMIQPGIYDEGRGRTWVW--VRFNPNSREWWARASGSRSP 57

DR 1 MASASVRLGPGRLGMIQPGIYDEGRGRTWVW--VRFNPNSREWWARASGSRSP 57

Qy 59 SIVVHLWQMAVHTRELLSSQMPFSQPLWVQKPKRAASSSFHEDLPHQDSDMQLVWYQPE 117

Db 58 S-----HLT-----SQLPLMMWOLYPEERYMDNNNSRLWQIQLHHMLVRYQVQELILKLPD 106

Qy 118 LVVYQPE 125

Db 99 LLLKLKLPD 106

Query Match 25.8%; Score 176.5; DB 1; Length 107;

Best Local Similarity 33.6%; Pred. No. 5.2e-12;

Matches 43; Conservative 17; Mismatches 43; Indels 25; Gaps 4;

DR 1 MASASVRLGPGRLGMIQPGIYDEGRGRTWVW--VRFNPNSREWWARASGSRSP 57

DR 1 MASASVRLGPGRLGMIQPGIYDEGRGRTWVW--VRFNPNSREWWARASGSRSP 57

Qy 59 SIVVHLWQMAVHTRELLSSQMPFSQPLWVQKPKRAASSSFHEDLPHQDSDMQLVWYQPE 117

Db 58 S-----HLT-----SQLPLMMWOLYPEERYMDNNNSRLWQIQLHHMLVRYQVQELILKLPD 106

Qy 118 LVVYQPE 125

Db 99 LLLKLKLPD 106

Query Match 25.8%; Score 176.5; DB 1; Length 107;

Best Local Similarity 33.6%; Pred. No. 5.2e-12;

Matches 43; Conservative 17; Mismatches 43; Indels 25; Gaps 4;

DR 1 MASASVRLGPGRLGMIQPGIYDEGRGRTWVW--VRFNPNSREWWARASGSRSP 57

DR 1 MASASVRLGPGRLGMIQPGIYDEGRGRTWVW--VRFNPNSREWWARASGSRSP 57

Qy 59 SIVVHLWQMAVHTRELLSSQMPFSQPLWVQKPKRAASSSFHEDLPHQDSDMQLVWYQPE 117

Db 58 S-----HLT-----SQLPLMMWOLYPEERYMDNNNSRLWQIQLHHMLVRYQVQELILKLPD 106

Qy 118 LVVYQPE 125

Db 99 LLLKLKLPD 106

Query Match 25.8%; Score 176.5; DB 1; Length 107;

Best Local Similarity 33.6%; Pred. No. 5.2e-12;

Matches 43; Conservative 17; Mismatches 43; Indels 25; Gaps 4;

DR 1 MASASVRLGPGRLGMIQPGIYDEGRGRTWVW--VRFNPNSREWWARASGSRSP 57

DR 1 MASASVRLGPGRLGMIQPGIYDEGRGRTWVW--VRFNPNSREWWARASGSRSP 57

Qy 59 SIVVHLWQMAVHTRELLSSQMPFSQPLWVQKPKRAASSSFHEDLPHQDSDMQLVWYQPE 117

Db 58 S-----HLT-----SQLPLMMWOLYPEERYMDNNNSRLWQIQLHHMLVRYQVQELILKLPD 106

Qy 118 LVVYQPE 125

Db 99 LLLKLKLPD 106

Query Match 25.8%; Score 176.5; DB 1; Length 107;

Best Local Similarity 33.6%; Pred. No. 5.2e-12;

Matches 43; Conservative 17; Mismatches 43; Indels 25; Gaps 4;

DR 1 MASASVRLGPGRLGMIQPGIYDEGRGRTWVW--VRFNPNSREWWARASGSRSP 57

DR 1 MASASVRLGPGRLGMIQPGIYDEGRGRTWVW--VRFNPNSREWWARASGSRSP 57

Qy 59 SIVVHLWQMAVHTRELLSSQMPFSQPLWVQKPKRAASSSFHEDLPHQDSDMQLVWYQPE 117

Db 58 S-----HLT-----SQLPLMMWOLYPEERYMDNNNSRLWQIQLHHMLVRYQVQELILKLPD 106

Qy 118 LVVYQPE 125

Db 99 LLLKLKLPD 106

Query Match 25.8%; Score 176.5; DB 1; Length 107;

Best Local Similarity 33.6%; Pred. No. 5.2e-12;

Matches 43; Conservative 17; Mismatches 43; Indels 25; Gaps 4;

DR 1 MASASVRLGPGRLGMIQPGIYDEGRGRTWVW--VRFNPNSREWWARASGSRSP 57

DR 1 MASASVRLGPGRLGMIQPGIYDEGRGRTWVW--VRFNPNSREWWARASGSRSP 57

Qy 59 SIVVHLWQMAVHTRELLSSQMPFSQPLWVQKPKRAASSSFHEDLPHQDSDMQLVWYQPE 117

Db 58 S-----HLT-----SQLPLMMWOLYPEERYMDNNNSRLWQIQLHHMLVRYQVQELILKLPD 106

Qy 118 LVVYQPE 125

Db 99 LLLKLKLPD 106

Query Match 25.8%; Score 176.5; DB 1; Length 107;

Best Local Similarity 33.6%; Pred. No. 5.2e-12;

Matches 43; Conservative 17; Mismatches 43; Indels 25; Gaps 4;

DR 1 MASASVRLGPGRLGMIQPGIYDEGRGRTWVW--VRFNPNSREWWARASGSRSP 57

DR 1 MASASVRLGPGRLGMIQPGIYDEGRGRTWVW--VRFNPNSREWWARASGSRSP 57

Qy 59 SIVVHLWQMAVHTRELLSSQMPFSQPLWVQKPKRAASSSFHEDLPHQDSDMQLVWYQPE 117

Db 58 S-----HLT-----SQLPLMMWOLYPEERYMDNNNSRLWQIQLHHMLVRYQVQELILKLPD 106

Qy 118 LVVYQPE 125

Db 99 LLLKLKLPD 106

Query Match 25.8%; Score 176.5; DB 1; Length 107;

Best Local Similarity 33.6%; Pred. No. 5.2e-12;

Matches 43; Conservative 17; Mismatches 43; Indels 25; Gaps 4;

DR 1 MASASVRLGPGRLGMIQPGIYDEGRGRTWVW--VRFNPNSREWWARASGSRSP 57

DR 1 MASASVRLGPGRLGMIQPGIYDEGRGRTWVW--VRFNPNSREWWARASGSRSP 57

Qy 59 SIVVHLWQMAVHTRELLSSQMPFSQPLWVQKPKRAASSSFHEDLPHQDSDMQLVWYQPE 117

Db 58 S-----HLT-----SQLPLMMWOLYPEERYMDNNNSRLWQIQLHHMLVRYQVQELILKLPD 106

Qy 118 LVVYQPE 125

Db 99 LLLKLKLPD 106

Query Match 25.8%; Score 176.5; DB 1; Length 107;

Best Local Similarity 33.6%; Pred. No. 5.2e-12;

Matches 43; Conservative 17; Mismatches 43; Indels 25; Gaps 4;

DR 1 MASASVRLGPGRLGMIQPGIYDEGRGRTWVW--VRFNPNSREWWARASGSRSP 57

DR 1 MASASVRLGPGRLGMIQPGIYDEGRGRTWVW--VRFNPNSREWWARASGSRSP 57

Qy 59 SIVVHLWQMAVHTRELLSSQMPFSQPLWVQKPKRAASSSFHEDLPHQDSDMQLVWYQPE 117

Db 58 S-----HLT-----SQLPLMMWOLYPEERYMDNNNSRLWQIQLHHMLVRYQVQELILKLPD 106

Qy 118 LVVYQPE 125

Db 99 LLLKLKLPD 106

Query Match 25.8%; Score 176.5; DB 1; Length 107;

Best Local Similarity 33.6%; Pred. No. 5.2e-12;

Matches 43; Conservative 17; Mismatches 43; Indels 25; Gaps 4;

DR 1 MASASVRLGPGRLGMIQPGIYDEGRGRTWVW--VRFNPNSREWWARASGSRSP 57

DR 1 MASASVRLGPGRLGMIQPGIYDEGRGRTWVW--VRFNPNSREWWARASGSRSP 57

Qy 59 SIVVHLWQMAVHTRELLSSQMPFSQPLWVQKPKRAASSSFHEDLPHQDSDMQLVWYQPE 117

Db 58 S-----HLT-----SQLPLMMWOLYPEERYMDNNNSRLWQIQLHHMLVRYQVQELILKLPD 106

Qy 118 LVVYQPE 125

Db 99 LLLKLKLPD 106

Query Match 25.8%; Score 176.5; DB 1; Length 107;

Best Local Similarity 33.6%; Pred. No. 5.2e-12;

Matches 43; Conservative 17; Mismatches 43; Indels 25; Gaps 4;

DR 1 MASASVRLGPGRLGMIQPGIYDEGRGRTWVW--VRFNPNSREWWARASGSRSP 57

DR 1 MASASVRLGPGRLGMIQPGIYDEGRGRTWVW--VRFNPNSREWWARASGSRSP 57

Qy 59 SIVVHLWQMAVHTRELLSSQMPFSQPLWVQKPKRAASSSFHEDLPHQDSDMQLVWYQPE 117

Db 58 S-----HLT-----SQLPLMMWOLYPEERYMDNNNSRLWQIQLHHMLVRYQVQELILKLPD 106

Qy 118 LVVYQPE 125

Db 99 LLLKLKLPD 106

Query Match 25.8%; Score 176.5; DB 1; Length 107;

Best Local Similarity 33.6%; Pred. No. 5.2e-12;

Matches 43; Conservative 17; Mismatches 43; Indels 25; Gaps 4;

DR 1 MASASVRLGPGRLGMIQPGIYDEGRGRTWVW--VRFNPNSREWWARASGSRSP 57

DR 1 MASASVRLGPGRLGMIQPGIYDEGRGRTWVW--VRFNPNSREWWARASGSRSP 57

Qy 59 SIVVHLWQMAVHTRELLSSQMPFSQPLWVQKPKRAASSSFHEDLPHQDSDMQLVWYQPE 117

Db 58 S-----HLT-----SQLPLMMWOLYPEERYMDNNNSRLWQIQLHHMLVRYQVQELILKLPD 106

Qy 118 LVVYQPE 125

Db 99 LLLKLKLPD 106

Query Match 25.8%; Score 176.5; DB 1; Length 107;

Best Local Similarity 33.6%; Pred. No. 5.2e-12;

Matches 43; Conservative 17; Mismatches 43; Indels 25; Gaps 4;

DR 1 MASASVRLGPGRLGMIQPGIYDEGRGRTWVW--VRFNPNSREWWARASGSRSP 57

DR 1 MASASVRLGPGRLGMIQPGIYDEGRGRTWVW--VRFNPNSREWWARASGSRSP 57

Qy 59 SIVVHLWQMAVHTRELLSSQMPFSQPLWVQKPKRAASSSFHEDLPHQDSDMQLVWYQPE 117

Db 58 S-----HLT-----SQLPLMMWOLYPEERYMDNNNSRLWQIQLHHMLVRYQVQELILKLPD 106

Qy 118 LVVYQPE 125

Db 99 LLLKLKLPD 106

Query Match 25.8%; Score 176.5; DB 1; Length 107;

Best Local Similarity 33.6%; Pred. No. 5.2e-12;

Matches 43; Conservative 17; Mismatches 43; Indels 25; Gaps 4;

DR 1 MASASVRLGPGRLGMIQPGIYDEGRGRTWVW--VRFNPNSREWWARASGSRSP 57

DR 1 MASASVRLGPGRLGMIQPGIYDEGRGRTWVW--VRFNPNSREWWARASGSRSP 57

Qy 59 SIVVHLWQMAVHTRELLSSQMPFSQPLWVQKPKRAASSSFHEDLPHQDSDMQLVWYQPE 117

Db 58 S-----HLT-----SQLPLMMWOLYPEERYMDNNNSRLWQIQLHHMLVRYQVQELILKLPD 106

Qy 118 LVVYQPE 125

Db 99 LLLKLKLPD 106

Query Match 25.8%; Score 176.5; DB 1; Length 107;

Best Local Similarity 33.6%; Pred. No. 5.2e-12;

Matches 43; Conservative 17; Mismatches 43; Indels 25; Gaps 4;

DR 1 MASASVRLGPGRLGMIQPGIYDEGRGRTWVW--VRFNPNSREWWARASGSRSP 57

DR 1 MASASVRLGPGRLGMIQPGIYDEGRGRTWVW--VRFNPNSREWWARASGSRSP 57

Qy 59 SIVVHLWQMAVHTRELLSSQMPFSQPLWVQKPKRAASSSFHEDLPHQDSDMQLVWYQPE 117

Db 58 S-----HLT-----SQLPLMMWOLYPEERYMDNNNSRLWQIQLHHMLVRYQVQELILKLPD 106

Qy 118 LVVYQPE 125

Db 99 LLLKLKLPD 106

Query Match 25.8%; Score 176.5; DB 1; Length 107;

Best Local Similarity 33.6%; Pred. No. 5.2e-12;

Matches 43; Conservative 17; Mismatches 43; Indels 25; Gaps 4;

DR 1 MASASVRLGPGRLGMIQPGIYDEGRGRTWVW--VRFNPNSREWWARASGSRSP 57

DR 1 MASASVRLGPGRLGMIQPGIYDEGRGRTWVW--VRFNPNSREWWARASGSRSP 57

Qy 59 SIVVHLWQMAVHTRELLSSQMPFSQPLWVQKPKRAASSSFHEDLPHQDSDMQLVWYQPE 117

Db 58 S-----HLT-----SQLPLMMWOLYPEERYMDNNNSRLWQIQLHHMLVRYQVQELILKLPD 106

Qy 118 LVVYQPE 125

Db 99 LLLKLKLPD 106

Query Match 25.8%; Score 176.5; DB 1; Length 107;

Best Local Similarity 33.6%; Pred. No. 5.2e-12;

Matches 43; Conservative 17; Mismatches 43; Indels 25; Gaps 4;

DR 1 MASASVRLGPGRLGMIQPGIYDEGRGRTWVW--VRFNPNSREWWARASGSRSP 57

DR 1 MASASVRLGPGRLGMIQPGIYDEGRGRTWVW--VRFNPNSREWWARASGSRSP 57

Qy 59 SIVVHLWQMAVHTRELLSSQMPFSQPLWVQKPKRAASSSFHEDLPHQDSDMQLVWYQPE 117

Db 58 S-----HLT-----SQLPLMMWOLYPEERYMDNNNSRLWQIQLHHMLVRYQVQELILKLPD 106

Qy 118 LVVYQPE 125

Db 99 LLLKLKLPD 106

Query Match 25.8%; Score 176.5; DB 1; Length 107;

Best Local Similarity 33.6%; Pred. No. 5.2e-12;

Matches 43; Conservative 17; Mismatches 43; Indels 25; Gaps 4;

DR 1 MASASVRLGPGRLGMIQPGIYDEGRGRTWVW--VRFNPNSREWWARASGSRSP 57

DR 1 MASASVRLGPGRLGMIQPGIYDEGRGRTWVW--VRFNPNSREWWARASGSRSP 57

Qy 59 SIVVHLWQMAVHTRELLSSQMPFSQPLWVQKPKRAASSSFHEDLPHQDSDMQLVWYQPE 117

Db 58 S-----HLT-----SQLPLMMWOLYPEERYMDNNNSRLWQIQLHHMLVRYQVQELILKLPD 106

Qy 118 LVVYQPE 125

Db 99 LLLKLKLPD 106

Query Match 25.8%; Score 176.5; DB 1; Length 107;

Best Local Similarity 33.6%; Pred. No. 5.2e-12;

Matches 43; Conservative 17; Mismatches 43; Indels 25; Gaps 4;

DR 1 MASASVRLGPGRLGMIQPGIYDEGRGRTWVW--VRFNPNSREWWARASGSRSP 57

DR 1 MASASVRLGPGRLGMIQPGIYDEGRGRTWVW--VRFNPNSREWWARASGSRSP 57

Qy 59 SIVVHLWQMAVHTRELLSSQMPFSQPLWVQKPKRAASSSFHEDLPHQDSDMQLVWYQPE 117

Db 58 S-----HLT-----SQLPLMMWOLYPEERYMDNNNSRLWQIQLHHMLVRYQVQELILKLPD 106

Q

**KW** Multigene family.  
**SEQUENCE** 122 AA; 14170 MW; BD1501F81C24F230 CRC64;  
**SO**

**Query Match** 25.6%; Score 175.5; DB 1; Length 122;  
**Best Local Similarity** 36.7%; Pred. No. 7.8e-12; Mismatches 50; Indels 7; Gaps 1;  
**Matches** 40; Conservative 12; Mismatches 50; Indels 7; Gaps 1;

**Qy** 12 PRGRMIQRCGTEDEBEGRTWTVVRENPSREWARASOGSRVPESTVHIMQAVHRR 71  
**Db** 13 PRFLVCTRDIDYBDENGRQWVAVV-----ETRSRPGSRKETCIVHILQHMTTQ 65  
**Qy** 72 ELLSGQMPQSQPAWVOLYGRKVRRAADSWEIADHGQDSMEQVLT 120  
**Db** 66 EPTPOQPNINNSLPTMWRSLMNTYTGTDGTVWRLLDHSQMDTQLQIL 114

**RESULT** 7  
**TCL1\_MOUSE** STANDARD; PRT; 116 AA.  
**ID** TCL1\_MOUSE  
**AC** P56280;  
**DT** 15-JUL-1998 (Rel. 36, Created)  
**DT** 30-MAY-2000 (Rel. 39, Last sequence update)  
**DE** T-CELL LEUKEMIA/LYMPHOMA PROTEIN 1A (PL4 TCL1 PROTEIN) (TCL1 ONCOGENE)  
**GN** (TCL1 OR TCL1).  
**OS** Mus musculus (Mouse).  
**OC** Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
**OC** Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
**OX** NCBI\_TaxID=10090;  
**RN** [1]  
**RP** SEQUENCE FROM N.A.  
**RX** MEDLINE=20056229; PubMed=10588720;  
**RA** Hallas C., Pekarsky Y., Itoyama T., Varnum J., Bichi R.,  
**RA** Rothstein J.I., Croce C.M.;  
**RT** Genomic analysis of human and mouse TCL1 loci reveals a complex of tightly clustered genes.;"  
**RL** Proc. Natl. Acad. Sci. U.S.A. 96:14418-14423 (1999).  
**CC** -!- SIMILARITY: BELONGS TO THE TCL1 FAMILY.  

**CC** This SWISS-PROT entry is copyright. It is produced through a collaboration between the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

**CC** EMBL: AF195493; AAF12806.1; -.  
**DR** MGI: MGI-1301635; Tc115.  
**DR** InterPro: IPR002709; Tc11\_MTCPI.  
**DR** ProDom: PD05575; Tc11\_MTCPI; 1.  
**KW** Multigene family.  
**SEQUENCE** 121 AA; 13668 MW; 32815CCD629B4297 CRC64;

**Query Match** 23.9%; Score 163.5; DB 1; Length 121;  
**Best Local Similarity** 33.0%; Pred. No. 1.4e-10; Mismatches 37; Conservative 14; Mismatches 44; Indels 17; Gaps 2;  
**Matches** 37; Conservative 14; Mismatches 44; Indels 17; Gaps 2;

**Qy** 19 ORP-----GYIYDEBEGRTWTVVRENPSREWARASOGSRVPESTVHIMQAVHRR 68  
**Db** 9 QRPFPVIVSVSLGTYEDHHRWIAVNV-----EDSHSHGURIERCIVHILQHMTT 61

**Qy** 69 HTRELLSGQMPQSQPAWVOLYGRKVRRAADSWEIADHGQDSMEQVLT 120  
**Db** 62 LPQPTPOQPNINNSLPTMWRSLMNTYTGTDGTVWRLLDHSQMDTQLQIL 113

**RESULT** 9  
**TCL2\_MOUSE** STANDARD; PRT; 117 AA.  
**ID** TCL2\_MOUSE  
**AC** P56841;  
**DT** 30-MAY-2000 (Rel. 39, Created)  
**DT** 20-AUG-2001 (Rel. 40, Last annotation update)  
**DE** TCL1B2 PROTEIN.  

**GN** TCL1B2.  
**OS** Mus musculus (Mouse).  
**OC** Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
**OC** Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
**OX** NCBI\_TaxID=10090;  
**RN** [1]  
**RP** SEQUENCE FROM N.A.  
**RX** MEDLINE=20056259; PubMed=10588720;  
**RA** Hallas C., Pekarsky Y., Itoyama T., Varnum J., Bichi R.,

**SEQUENCE** 116 AA; 14112 MW; 46DDEED2F973F389A CRC64;

**Query Match** 24.7%; Score 169; DB 1; Length 116;  
**Best Local Similarity** 32.8%; Pred. No. 3.8e-11; Mismatches 40; Conservative 20; Mismatches 44; Indels 18; Gaps 2;

**Matches** 40; Conservative 20; Mismatches 44; Indels 18; Gaps 2;

**Qy** 1 MASERSVYRLGVPPGRMLTQPGGIVDEBEGRTWTVVRENPSREWARASOGSRVPEPS 58  
**Db** 1 MATORRAHAEPAHNRWIKERHHLDEFRRSWLPVWIKSN-----EK 44

**Qy** 59 ITWHLWQMAVHTRRELSSQMPFSQPAWVOLYGRKVRRAADSWEIADHGQDSMEQVLT 118  
**Db** 45 FQVILRQEDVTVGAEAMSPSQLVPEFLPLMQLYPLQKDRYRSQDSMWMQIUYHIFROVEMD 104

**Qy** 119 VL 120





DT 01-JAN-1990 (Rel. 1.3; Last sequence update) RT 20-AUG-2001 (Rel. 4.0; Last annotation update) RT T-DNA BORDER ENDONUCLEASE VIRD2 (EC 3.1.1.-) RN OS Agrobacterium rhizogenes. RN OC Plasmid PR1Ab. RN OC Bacterium; Proteobacteria; alpha subdivision; Rhizobiaceae group; RN OC Rhizobiaceae; Rhizobium. RN NCBI\_TAXID=559; RN [1] SEQUENCE FROM N.A. RN RP MEDLINE=89039712; PubMed=3185501; RN RA Hirayama T., Muranaka T., Okawa H., Oka A.; RN RT "Organization and characterization of the virCD genes from Agrobacterium rhizogenes."; RT Mol. Genet. 23:229-237 (1988) RT -!- FUNCTION: TUMOR FORMATION BY A TUMEFACIENS INVOLVES THE TRANSFER AND INTEGRATION OF A DEFINED SEGMENT (T-DNA) OF TI PLASMID DNA INTO THE PLANT NUCLEAR GENOME. THE VIRD OPERON ENCODES A SITE-SPECIFIC ENDONUCLEASE THAT CLEAVES AT A UNIQUE SITE WITHIN BOTH 24 BP DIRECT REPEATS FLANKING THE T-DNA. RT -----

This SWISS-PROT entry is copyright. It is produced through collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; X12847; CA31351.1; -. CC DR PIR; SU06884; S06884. CC KW Hydrolase; Nuclease; Endonuclease; Crown gall tumor; Plasmid; T-DNA. CC SQ SEQUENCE 436 AA; 48371 MW; 932A465E0966A6 CRC64;

Query Match 10.9%; Score 74.5; DB 1; Length 436; Best Local Similarity 25.6%; Pred. No. 1.7; Matches 34; Conservative 17; Mismatches 43; Indels 39; Gaps 8; QY 11 VPP-----GRMIORGTE---DESGERTWY---VVVER-----NPSREWARA 47 Db 49 VPPDQIHELARSWVQEGTIDESQPDERQQLTITHIIVSFPAGTSQAAAYASREWARAE 108 QY 48 --RASQGSRVEPSTVHL-----WOMAVHRELLSSGOMPSPQLPAPWLYPERKRYA 98 Db 109 MFGSGAGGSYVNLTAFHIDRHIFHLHVWNRRELLGHGNLKS-----RHPQUNYDA 162 QY 99 ADSSTWETA-DHG 110 Db 163 IRINMAEISLRHG 175

RESULT 14

VIRD2\_AGRT6 STANDARD; PRT; 424 AA.

Query Match 10.6%; Score 72.5; DB 1; Length 424; Best Local Similarity 24.1%; Pred. No. 2.7; Matches 33; Conservative 24; Mismatches 49; Indels 47; Gaps 8; QY 11 VPPRL-----WIPORGTE---DEBGRGWWY---VVYRF-----NPSREWARA 49 Db 49 VPPDQIHELARSWVQEGTIDESQPDERQQLTITHIIVSFPAGTQDTAAYESREWARAE 108 QY 50 SGS-----RYPSTVHL-----WOMAVHRELLSSGOMPSPQLPAPWLYPERKRYA 97 Db 109 MFGSGAGGSYVNLTAFHIDRHIFHLHVWNRRELLGHGNLKS-----RHPQUNYDA 166 QY 98 AADSFNWITAHGQI-----DSMQLVITYPER 127 Db 167 MAEISL-----RHGIVLDATSAERGIAERPYAEHRR 200

RESULT 15

GIGB\_S01TU STANDARD; PRT; 861 AA.

Query Match 10.6%; Score 72.5; DB 1; Length 424; Best Local Similarity 24.1%; Pred. No. 2.7; Matches 33; Conservative 24; Mismatches 49; Indels 47; Gaps 8; QY 11 VPPRL-----WIPORGTE---DEBGRGWWY---VVYRF-----NPSREWARA 49 Db 49 VPPDQIHELARSWVQEGTIDESQPDERQQLTITHIIVSFPAGTQDTAAYESREWARAE 108 QY 50 SGS-----RYPSTVHL-----WOMAVHRELLSSGOMPSPQLPAPWLYPERKRYA 97 Db 109 MFGSGAGGSYVNLTAFHIDRHIFHLHVWNRRELLGHGNLKS-----RHPQUNYDA 166 QY 98 AADSFNWITAHGQI-----DSMQLVITYPER 127 Db 167 MAEISL-----RHGIVLDATSAERGIAERPYAEHRR 200

RESULT 16

GLGB\_S01TU STANDARD; PRT; 861 AA.

Query Match 10.6%; Score 72.5; DB 1; Length 424; Best Local Similarity 24.1%; Pred. No. 2.7; Matches 33; Conservative 24; Mismatches 49; Indels 47; Gaps 8; QY 11 VPPRL-----WIPORGTE---DEBGRGWWY---VVYRF-----NPSREWARA 49 Db 49 VPPDQIHELARSWVQEGTIDESQPDERQQLTITHIIVSFPAGTQDTAAYESREWARAE 108 QY 50 SGS-----RYPSTVHL-----WOMAVHRELLSSGOMPSPQLPAPWLYPERKRYA 97 Db 109 MFGSGAGGSYVNLTAFHIDRHIFHLHVWNRRELLGHGNLKS-----RHPQUNYDA 166 QY 98 AADSFNWITAHGQI-----DSMQLVITYPER 127 Db 167 MAEISL-----RHGIVLDATSAERGIAERPYAEHRR 200

RESULT 17

SEQUENCE FROM N.A. MEDLINE=87038239; PubMed=3021341; Yanocho M.F.; Porter S.G.; Young C.; Albright L.M.; Gordon M.P.; Nester E.W.;

RT "The virD operon of Agrobacterium tumefaciens encodes a site-specific endonuclease."; RT Cell 47:471-477 (1986). RT [2] SEQUENCE FROM N.A. RT MEDLINE=88032822; PubMed=2822660; RT RA Jayaswal R.K.; Veiluthambi K.; Gelvin S.B.; Slightom J.L.; RT "Double-stranded cleavage of T-DNA and generation of single-stranded T-DNA molecules in *Escherichia coli* by a virD-encoded border-specific endonuclease from Agrobacterium tumefaciens."; RT RLU J. Bacteriol. 169:5035-5045 (1987). RT CC AND INTEGRATION OF A DEFINED SEGMENT (T-DNA) OF TI PLASMID DNA INTO THE PLANT NUCLEAR GENOME. THE VIRD OPERON ENCODES A SITE-SPECIFIC ENDONUCLEASE THAT CLEAVES AT A UNIQUE SITE WITHIN BOTH 24 BP DIRECT REPEATS FLANKING THE T-DNA. RT -----

This SWISS-PROT entry is copyright. It is produced through collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; AF242881; AAA98390.1; -. CC DR EMBL; ML1989; AAA22114.1; -. CC PIR; B29826; B29826. CC DR PIR; B25063; B25063. CC KW Hydrolyase; Nuclease; Endonuclease; Crown gall tumor; Plasmid; T-DNA. CC SQ SEQUENCE 424 AA; 47546 MW; B2CB7B82C01A6C1 CRC64;

Query Match 10.6%; Score 72.5; DB 1; Length 424; Best Local Similarity 24.1%; Pred. No. 2.7; Matches 33; Conservative 24; Mismatches 49; Indels 47; Gaps 8; QY 11 VPPRL-----WIPORGTE---DEBGRGWWY---VVYRF-----NPSREWARA 49 Db 49 VPPDQIHELARSWVQEGTIDESQPDERQQLTITHIIVSFPAGTQDTAAYESREWARAE 108 QY 50 SGS-----RYPSTVHL-----WOMAVHRELLSSGOMPSPQLPAPWLYPERKRYA 97 Db 109 MFGSGAGGSYVNLTAFHIDRHIFHLHVWNRRELLGHGNLKS-----RHPQUNYDA 166 QY 98 AADSFNWITAHGQI-----DSMQLVITYPER 127 Db 167 MAEISL-----RHGIVLDATSAERGIAERPYAEHRR 200

RESULT 18

SEQUENCE FROM N.A. MEDLINE=9410534; PubMed=8278528; Poulsen P.; Krasberg J.D.; "Starch branching enzyme, cDNA from *Solanum tuberosum*;" plant physiol. 100:103-107 (1992). RT

RN	[2]
RP	SEQUENCE OF 279-527 FROM N-A.
RC	STRAIN=CV; DESEREE; TISSUE=Tuber;
RX	MEDLINE=92079917; PubMed=145241;
RA	Kossmann J., Visser R.G.F., Mueller-Roeber B., Wilmitzter L., Sonnewald U.;
RA	"Cloning and expression analysis of a potato cDNA that encodes a branching enzyme: evidence for co-expression of starch biosynthetic genes.", Mol. Genet. 230:39-44 (1991).
RT	-1- CATALYTIC ACTIVITY: FORMATION OF 1,6-GLUCOSIDIC LINKAGES OF STARCH.
CC	-1- PATHWAY: THIRD STEP IN STARCH BIOSYNTHESIS.
CC	-1- SUBUNIT: MONOMER.
CC	-1- SUBCELLULAR LOCATION: AMYLOPLAST.
CC	-1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).
CC	-----
CC	-----
DR	EMBL: X69805; CAA49463.1; -.
DR	PIR: S18394; S18394.
DR	Mendel: 13376; S01tu; Sbel; 1.
DR	InterPro: IPR000461; Alpha_amylase.
DR	PFam: PF00128; alpha_amylase_1.
KW	Starch biosynthesis; Transferase; Glycosyltransferase; Amyloplast.
FT	ACT_SITE 424 424 BY SIMILARITY.
FT	ACT_SITE 484 484 BY SIMILARITY.
FT	ACT_SITE 553 553 BY SIMILARITY.
FT	ACT_SITE 861 861 AA; 99083 MW; F3D519AC7CF1BEE2 CRC64; SEQUENCE
Query	Match
Best	Loc1 Similarity
Matches	24.28; Pred. No. 11; Length 861;
31;	Conservative
OY	17 WIQRGIVDEEERTWTVWVRFNSRREWA-SGDSRVSPTVHQMVAWHTR-- 72
Db	620 WIDFP-----REGNNW-----SYDKCRROWNLADSEHLRYK-- FMAFDRAANSLEKF 666
OY	73 -LSSQGMFQ-----OLPAVWQLYPGR-----KYNAA-DSSF 103
Db	667 SFLASGKQTVSSMDDNNKVVVFERGDLVVFVNHFHKPKNTYEGYKVGCDLPGKRYVALDSA 726
OY	104 WRIADHQ 111
Db	727 WFGGIGR 734

Search completed: November 29, 2001, 04:06:42  
Job time: 208 sec

Gencore version 4.5  
copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using SW model

Run on : November 29, 2001, 04:03:39 ; Search time 23.95 Seconds  
(without alignments) 781.748 Million cell updates/sec

Title: US-09-526-329-39  
Perfect score: 685

Sequence:

1 MASEASVRLGVPPGRLLIOPRGPYDEBGRWVWVWVNPNSREWARASQGSRYEPIT

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched:

473505 seqs, 14627329 residues

Total number of hits satisfying chosen parameters:

473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL\_17:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rhodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

### SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	677	98.8	128	4 Q9UBQ4
2	113	16.5	107	11 Q9UXN9
3	81	11.8	448	2 Q9FB87
4	79.5	11.6	447	2 Q44461
5	79.5	11.6	447	2 Q9RB9
6	79	11.5	238	2 Q9HZ1
7	76.5	11.2	588	12 Q70681
8	76	11.1	3742	8 Q9TRW8
9	75.5	11.0	476	2 Q9KDN8
10	75	10.9	986	4 Q9UFO
11	73	10.7	463	2 Q9XK7
12	73	10.7	464	10 Q9FT19
13	73	10.7	513	2 Q918Q3
14	73	10.7	2229	2 Q9KRP5
15	72.5	10.6	513	2 Q9B939
16	72.5	10.6	1414	2 P7117
17	70	10.2	367	2 Q9KV7
18	70	10.2	477	12 Q9ILY2
19	70	10.2	830	10 Q04864

Database :

SPREMBL\_17:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rhodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*

Total number of hits satisfying chosen parameters:

473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL\_17:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rhodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*

Total number of hits satisfying chosen parameters:

473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL\_17:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rhodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*

Total number of hits satisfying chosen parameters:

473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL\_17:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rhodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*

Total number of hits satisfying chosen parameters:

473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL\_17:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rhodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*

Total number of hits satisfying chosen parameters:

473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL\_17:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rhodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*

Total number of hits satisfying chosen parameters:

473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL\_17:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rhodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*

Total number of hits satisfying chosen parameters:

473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL\_17:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rhodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*

Total number of hits satisfying chosen parameters:

473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL\_17:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rhodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*

Total number of hits satisfying chosen parameters:

473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL\_17:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rhodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*

Total number of hits satisfying chosen parameters:

473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL\_17:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rhodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*

Total number of hits satisfying chosen parameters:

473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL\_17:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rhodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*

Total number of hits satisfying chosen parameters:

473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL\_17:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rhodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*

Total number of hits satisfying chosen parameters:

473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL\_17:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rhodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*

Total number of hits satisfying chosen parameters:

473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL\_17:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rhodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*

Total number of hits satisfying chosen parameters:

473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL\_17:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rhodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*

Total number of hits satisfying chosen parameters:

QY	121	TYQPERKD	128	[ 3 ]
RP		SEQUENCE FROM N.A.		
RC		STRAIN=MAFF03-01724;		
RA		Moriguchi K., Nishida T., Maeda Y., Tanaka N., Yoshida K.;		
RT		"Genomic structure of Ri plasmid (1)-construction of linking library and physical map of pR1724 in Japanese Agrobacterium.;"		
RL		Nucleic Acids Symp. Ser. 39:189-190(1998).		
RN	[ 4 ]	SEQUENCE FROM N.A.		
RP		STRAIN=MAFF03-01724;		
RC		STRAIN=MAFF03-01724;		
RA		MEDLINE=201294; PubMed=10780382;		
RT		Maeda Y., Moriguchi K., Kataoka M., Satou M., Satuti N., Tanaka N., Yoshida K.;		
RT		"Genomic structure of Ri plasmid (1): Sequencing analysis of T-DNA and its flanking regions of pR1724 in Japanese Agrobacterium rhizogenes;"		
RT		Nucleic Acids Symp. Ser. 42:67-68(1999).		
RL		EMBL: AP002085; BAB16287.1; -;		
DR		Plasmid.		
KW		SEQUENCE FROM N.A.		
RX		SEQUENCE FROM N.A.		
RA		SEQUENCE FROM N.A.		
RT		SEQUENCE FROM N.A.		
RL		SEQUENCE FROM N.A.		
RN	[ 1 ]	SEQUENCE FROM N.A.		
RX		MEDLINE=20056255; PubMed=10588720;		
RA		Hallas C., Pekarsky Y., Itaya T., Varnum J., Bichl R., Rotstein J.I., Croce C.M.;		
RT		"Genomic analysis of human and mouse T-cell loci reveals a complex of tightly clustered genes.;"		
RL		Proc. Natl. Acad. Sci. U.S.A. 96:14418-14423(1999).		
DR		EMBL: AF195191; AACF2804.1; -;		
DR		MGD: MG1-1135160; TCI: B3.		
SO		SEQUENCE 107 AA; 12244 MW; DDC2AAC49D02DBE CRC64;		
SO		SEQUENCE 448 AA; 50021 MW; FB2594921BCE8704 CRC64;		
Query	Match	16.5%; Score 113; DB 11; Length 107;	11.8%; Score 81; DB 2; Length 448;	
Best Local Similarity	31.6%;	Pred. No. 0.00022;	Best Local Similarity 26.8%; Pred. No. 2.4;	
Matches	30;	Conservative 13; Mismatches 45; Indels 6; Gaps 3;	Matches 40; Conservative 16; Mismatches 49; Indels 44; Gaps	
QY	29	GRTAWTVVPRFSPREWARS---QSSRYFESVTIVHLWQMAWVHTRELLSQMPFSQLP 85	11 VPP-----GRLMQRCIVE---DESRGTVW-----VVWRP-----NPSRREWA-- 47	
Db	8	GLQWAS--LRFS-GHALORASIKLKSGLHLLTCITYVHLOHMTTIPQEPPTQPOQPNINNSLP 64	11 VPP-----GRLMQRCIVE---DESRGTVW-----VVWRP-----NPSRREWA-- 47	
QY	86	AVWQLYPGRKRYRADSSFWLADHGQDMSMQLVL 120	49 VPPDEIRELARSLWQETGTYHESQDDEERQDQLTHTIVSPAGTSQAYASREWAE 108	
Db	65	TMWRLDEASNTVYGTDTWMLDLSQGDTQLLIT 99	48 --RASGSKYFESVTIVL-----WQMAWVHTRELLSQMPFSQLPAWVWYGRKRY 98	
QY	95587	PRELIMINARY; PRT; 448 AA.	109 MFGSGAGGGRNLYTAFHIDRDPHLHVWNRKELLGHGLWIKS-----RHHPOLNDA 162	
ID	QPF587		Db 99 ADSSFWELA DHGQDMSQELVLTQPER 163 LRIKMSSETSLRHGVV-----LEATRRAER 186	
AC	OPF587;			
DT	01-MAR-2001 (TREMBREL. 16, created)			
DT	01-MAR-2001 (TREMBREL. 16, last sequence update)			
DT	01-MAR-2001 (TREMBREL. 16, last annotation update)			
DE	RIOF168 PROTEIN.			
DE	RIOF168 PROTEIN.			
GN	OS			
OS	Arrobacterium rhizogenes.			
OG	OG			
OC	Bacterium: Rhizobacterium; alpha subdivision; Rhizobiaceae group; Rhizobium.			
OC	Bacterium: Proteobacteria; alpha subdivision; Rhizobiaceae group; Rhizobium.			
OX	NCBI_TAXID=359;			
RN	[ 1 ]	SEQUENCE FROM N.A.		
RN	SEQUENCE FROM N.A.			
RN	SEQUENCE FROM N.A.			
RN	SEQUENCE FROM N.A.			
RP		Wang K., Herrera-Estrella A.H., Van Montagu M.M.; Submitted (JUL-1990) to the EMBL/GenBank/DBJ databases.		
RC		EMBL: M31673; AAA22111.1; -;		
RC		SEQUENCE 447 AA; 49649 MW; 9F433A5014B42256 CRC64;		
RC		DR		
RC		SEQUENCE FROM N.A.		
RC		STRAIN=MAFF03-01724;		
RC		Moriguchi K., Maeda Y., Satou M., Satuti N., Kataoka M., Tanaka N., Yoshida K.;		
RC		"The complete nucleotide sequence of a Ri (root inducing) plasmid indicates its chemical structure between Ti and Sym plasmids.;"		
RT		Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.		
RT		[ 2 ]		
RT		SEQUENCE FROM N.A.		
RT		STRAIN=MAFF03-01724;		
RT		Moriguchi K., Maeda Y., Satou M., Kataoka M., Tanaka N., Yoshida K.;		
RT		"Analysis of unique variable region of a plant root inducing plasmid, PR1724, by the construction of its physical map and library.;"		
RT		Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.		
RL		[ 1 ]		

Db	163 IRIKMAEISLRHG 175	Qy	48 --RASQGSRVERPSITVHL----WQAVHTRELLSGOMPSQLPAVWOLYPLGKVRKA 98
RESULT	5	Q9H7Z1	PRELIMINARY; PRT; 447 AA.
Q9R6B9	PRELIMINARY;	Q9H7Z1	PRELIMINARY; PRT; 238 AA.
AC	Q9R6B9;	Q9H7Z1;	AC
DT	01-MAY-2000 (TREMBLrel. 13, Created)	Q9H7Z1;	DT
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)	Q9H7Z1;	DT
DT	01-MAY-2001 (TREMBLrel. 16, Last annotation update)	Q9H7Z1;	DT
DE	T10R156 PROTEIN	Q9H7Z1	DE
GN	T10R156.	Q9H7Z1	GN
OS	Agrobacterium radiobacter.	Q9H7Z1	OS
OG	Plasmid pti-SAKURA.	Q9H7Z1	OG
OC	Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;	Q9H7Z1	OC
OC	Rhizobiaceae; Rhizobium.	Q9H7Z1	OC
OX	NCBI_TaxID=388;	Q9H7Z1	OX
RN	[1]	Q9H7Z1	RN
RP	SEQUENCE FROM N.A.	Q9H7Z1	RP
RC	STRAIN=MAFF301001; PubMed=10721727;	Q9H7Z1	RC
RX	MEDLINE=20184752;	Q9H7Z1	RX
RA	Suzuki K., Hattori Y., Uraji M., Ohta N., Iwata K., Murata K.,	Q9H7Z1	RA
RA	Katoh A., Yoshida K.;	Q9H7Z1	RA
RT	*Complete nucleotide sequence of a plant tumor-inducing Ti plasmid.;	Q9H7Z1	RT
RL	Gene 242:331-336(2000).	Q9H7Z1	RL
RN	[2]	Q9H7Z1	RN
RP	SEQUENCE FROM N.A.	Q9H7Z1	RP
RC	STRAIN=MAFF301001;	Q9H7Z1	RC
RX	MEDLINE=98193120; PubMed=9524202;	Q9H7Z1	RX
RA	Suzuki K., Ohta N., Hattori Y., Uraji M., Katoh A., Yoshida K.;	Q9H7Z1	RA
RT	*Novel structural difference between nopaline- and octopine-type trbJ gene-construction of genetic and physical map and sequencing of trbJ and rep gene clusters of a new Ti plasmid pti-SAKURA.;	Q9H7Z1	RT
RT	trbJ/trbI and rep gene clusters of a new Ti plasmid pti-SAKURA.;	Q9H7Z1	RT
RL	Biophys. Acta 1396:11-7(1998).	Q9H7Z1	RL
RN	[3]	Q9H7Z1	RN
RP	SEQUENCE FROM N.A.	Q9H7Z1	RP
RC	STRAIN=MAFF301001;	Q9H7Z1	RC
RA	Hattori Y., Suzuki K., Ohta N., Uraji M., Katoh A., Yoshida K.;	Q9H7Z1	RA
RT	*Genome structure of pti-SAKURA(I): Strategy for DNA sequencing of a Japanese cherry Ti plasmid.	Q9H7Z1	RT
RL	PLASMID.;	Q9H7Z1	RL
RA	Nucleic Acids Symp. Ser. 37:159-160(1998).	Q9H7Z1	RA
RN	[4]	Q9H7Z1	RN
RP	SEQUENCE FROM N.A.	Q9H7Z1	RP
RC	STRAIN=MAFF301001;	Q9H7Z1	RC
RA	Ohta N., Suzuki K., Hattori Y., Uraji M., Katoh A., Yoshida K.;	Q9H7Z1	RA
RT	*Genome structure of pti-SAKURA (II): Characteristics of T-DNA.;	Q9H7Z1	RT
RL	Nucleic Acids Symp. Ser. 39:185-186(1998).	Q9H7Z1	RL
RN	[5]	Q9H7Z1	RN
RP	SEQUENCE FROM N.A.	Q9H7Z1	RP
RC	STRAIN=MAFF301001;	Q9H7Z1	RC
RA	Uraji M., Suzuki K., Ohta N., Hattori Y., Katoh A., Yoshida K.;	Q9H7Z1	RA
RT	*Genome structure of pti-SAKURA (IV): Characteristics of tra region.;	Q9H7Z1	RT
RL	Nucleic Acids Symp. Ser. 39:187-188(1998).	Q9H7Z1	RL
RN	[6]	Q9H7Z1	RN
RP	SEQUENCE FROM N.A.	Q9H7Z1	RP
RC	STRAIN=MAFF301001;	Q9H7Z1	RC
RA	Hattori Y., Suzuki K., Ohta N., Uraji M., Katoh A., Yoshida K.;	Q9H7Z1	RA
RT	*Genome structure of pti-SAKURA (VI): Complete nucleotide sequence of plasmid pti-SAKURA vir region in Agrobacterium tumefaciens.;	Q9H7Z1	RT
RL	Nucleic Acids Symp. Ser. 39:265-266(1998).	Q9H7Z1	RL
DR	EMBL: AB016260; BAM87781.1; -	Q9H7Z1	DR
RW	Plasmid.	Q9H7Z1	RW
SO	SEQUENCE 447 AA; 49603 MW; 5A9B646F2DB8894AC CRC64;	Q9H7Z1	SO
Query Match	11.5%; Score 79.5; DB 2; Length 238;	Q9H7Z1	Query Match
Best Local Similarity	23.9%; Pred. No. 1.8; Gaps 3;	Q9H7Z1	Best Local Similarity
Matches	28; Conservative 17; Mismatches 46; Indels 26; Gaps 3;	Q9H7Z1	Matches
Qy	5 ASVRIGVPGPRWIMQRPGLYDEBEGRWWVWRNPSPREMASQGSRVERPSITVHLN 64	Q9H7Z1	Qy
Db	42 ATTRILREALQIOLQSGQLYREB-----RRGWFVSPERILAYNPLVSHEH 87	Q9H7Z1	Db
Qy	65 QMA---WHRLELLSGQMF---QLPAWQIIPGKRYRADSSFWIADH 109	Q9H7Z1	Qy
Db	88 AMVSEQGRVPAVEILSAROMPASAANVCELLPALSVYQIRRARVYDGLVLYEH 144	Q9H7Z1	Db
RESULT	7	Q9H7Z1	RESULT
OT0681	PRELIMINARY; PRT; 588 AA.	Q9H7Z1	OT0681
ID	070681	Q9H7Z1	ID
AC	070681;	Q9H7Z1	AC
DT	01-AUG-1998 (TREMBLrel. 07, Created)	Q9H7Z1	DT
DT	01-AUG-1998 (TREMBLrel. 07, Last sequence update)	Q9H7Z1	DT
DE	COAT PROTEIN (FRAGMENT)	Q9H7Z1	DE
OC	Sugarcane streak mosaic virus.	Q9H7Z1	OC
OC	Viruses: SRNA positive-strand viruses, no DNA stage; Potyviridae;	Q9H7Z1	OC
OC	Potyvirus.	Q9H7Z1	OC
NCBI_TaxID=53954;	Q9H7Z1	NCBI_TaxID=53954;	RN
[1]	Q9H7Z1	[1]	RN



OX	NCBL_TAXID=9606;	Query Match	10.7%	Score	73;	DB	2;	Length	463;		
RN	[1]	Best Local Similarity	24.3%	Pred.	No.	17;	Mismatches	35;	Indels	36;	Gaps
RP	SEQUENCE FROM N.A.	Matches	35;	Conservative	18;	Mismatches	55;	Indels	36;	Gaps	
RC	TISSUE-BRAIN;										
RX	MEDLINE=20039619; PubMed=10574462;	1	MASEASVRLGVPGRLWTOPRG----YDEDEGRWVVW-----VRF	39							
RA	Narase T., Ishikawa K., Kikuno R., Hirosva M., Nomura N., Ohara O.;	:		:							
RT	"Prediction of the coding sequences of unidentified human genes. XV. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro.;"										
RT	EMBL: AB032096; BA86584.1;										
DR	InterPro; IPR02106; AA_TRNA_ligase_II.										
DR	InterPro; IPR002318; tRNA-synt_2c.										
DR	PRINTS; PRO0980; TNAASINTHALA.										
DR	PROSITE; PS0039; AA_TRNA_LIGASE_II_2; UNKNOWN 1.										
FT	SEQUENCE	986 AA;	107427 MW;	E6FDD07907AR831	CRC64;						
FT	NON_TER	1									
SQ											
RESULT	11	Query Match	10.9%	Score	75;	DB	4;	Length	986;		
Q9HXM7	09HAM7	Best Local Similarity	25.0%	Pred.	No.	25;	Mismatches	51;	Indels	30;	Gaps
AC	Q9HAM7;	Matches	34;	Conservative	16;	Mismatches	51;	Indels	30;	Gaps	8;
QY	10 GYPPGRLWLGQ----REGIYDESGR-TWTVY----VVRPNPSRREWARASQSRVE 56	10	GYPPGRLWLGQ----REGIYDESGR-TWTVY----VVRPNPSRREWARASQSRVE 56	56							
Db	162 GIEPFERLWISYFDGDPKAGLDPDLTRDVLISLGVPASRVLISFGPQENEWMGDTGP-CG 220	162	GIEPFERLWISYFDGDPKAGLDPDLTRDVLISLGVPASRVLISFGPQENEWMGDTGP-CG 220	220							
QY	57 PSITVHLWOMAHVRELLSGQMPFSQALPAWQOLYPPKRYRAAASSFWEI----ADHGQI 112	57	PSITVHLWOMAHVRELLSGQMPFSQALPAWQOLYPPKRYRAAASSFWEI----ADHGQI 112	112							
Db	221 PCNEIH-YDKA-----GGVGAPOLYELWNLVFMHNRREADGSLQPLPORHWDGM- 269	221	PCNEIH-YDKA-----GGVGAPOLYELWNLVFMHNRREADGSLQPLPORHWDGM- 269	269							
QY	113 DSNEQLVLTQ 123	113	DSNEQLVLTQ 123	123							
Db	270 -GLERLWAVLQ 279	270	-GLERLWAVLQ 279	279							
RESULT	12	Query Match	10.9%	Score	75;	DB	4;	Length	986;		
Q91199	09F199	Best Local Similarity	25.0%	Pred.	No.	25;	Mismatches	51;	Indels	30;	Gaps
AC	09F199;	Matches	34;	Conservative	16;	Mismatches	51;	Indels	30;	Gaps	8;
QY	113 DSNEQLVLTQ 123	113	DSNEQLVLTQ 123	123							
Db	270 -GLERLWAVLQ 279	270	-GLERLWAVLQ 279	279							
RESULT	12	Query Match	10.7%	Score	73;	DB	10;	Length	464;		
Q91199	09F199	Best Local Similarity	23.8%	Pred.	No.	17;	Mismatches	31;	Indels	49;	Gaps
AC	09F199;	Matches	31;	Conservative	18;	Mismatches	49;	Indels	32;	Gaps	
QY	16 LWIWRPGIYDEEGRWT-----VVRPNPSRREWARASQSRVEPSTV 61	16	LWIWRPGIYDEEGRWT-----VVRPNPSRREWARASQSRVEPSTV 61	61							
Db	303 LWIWRPGIYDEEGRWT-----VVRPNPSRREWARASQSRVEPSTV 61	303	LWIWRPGIYDEEGRWT-----VVRPNPSRREWARASQSRVEPSTV 61	61							
QY	62 HILWOMAHVRELLSGQMPFSQALPAWQOLYPPKRYRAAASSFWEIADHGQI 117	62	HILWOMAHVRELLSGQMPFSQALPAWQOLYPPKRYRAAASSFWEIADHGQI 117	117							
Db	356 NGWN----STLIESICIG-VPMICLICKNDQFWNARE--ISEFWRVLGHLERREI 408	356	NGWN----STLIESICIG-VPMICLICKNDQFWNARE--ISEFWRVLGHLERREI 408	408							
QY	118 EYVLTQPERK 127	118	EYVLTQPERK 127	127							
Db	409 AVIRLWESK 418	409	AVIRLWESK 418	418							
RESULT	13	Query Match	10.7%	Score	73;	DB	2;	Length	463;		
Q918Q3	Q918Q3	Best Local Similarity	24.3%	Pred.	No.	17;	Mismatches	35;	Indels	36;	Gaps
AC	Q918Q3;	Matches	35;	Conservative	18;	Mismatches	55;	Indels	36;	Gaps	
QY	118 EYVLTQPERK 127	118	EYVLTQPERK 127	127							
Db	409 AVIRLWESK 418	409	AVIRLWESK 418	418							
RESULT	13	Query Match	10.7%	Score	73;	DB	2;	Length	463;		
Q918Q3	Q918Q3	Best Local Similarity	24.3%	Pred.	No.	17;	Mismatches	35;	Indels	36;	Gaps
AC	Q918Q3;	Matches	35;	Conservative	18;	Mismatches	55;	Indels	36;	Gaps	
QY	118 EYVLTQPERK 127	118	EYVLTQPERK 127	127							
Db	409 AVIRLWESK 418	409	AVIRLWESK 418	418							
RESULT	13	Query Match	10.7%	Score	73;	DB	2;	Length	463;		
Q918Q3	Q918Q3	Best Local Similarity	24.3%	Pred.	No.	17;	Mismatches	35;	Indels	36;	Gaps
AC	Q918Q3;	Matches	35;	Conservative	18;	Mismatches	55;	Indels	36;	Gaps	
QY	118 EYVLTQPERK 127	118	EYVLTQPERK 127	127							
Db	409 AVIRLWESK 418	409	AVIRLWESK 418	418							
RESULT	13	Query Match	10.7%	Score	73;	DB	2;	Length	463;		
Q918Q3	Q918Q3	Best Local Similarity	24.3%	Pred.	No.	17;	Mismatches	35;	Indels	36;	Gaps
AC	Q918Q3;	Matches	35;	Conservative	18;	Mismatches	55;	Indels	36;	Gaps	
QY	118 EYVLTQPERK 127	118	EYVLTQPERK 127	127							
Db	409 AVIRLWESK 418	409	AVIRLWESK 418	418							
RESULT	13	Query Match	10.7%	Score	73;	DB	2;	Length	463;		
Q918Q3	Q918Q3	Best Local Similarity	24.3%	Pred.	No.	17;	Mismatches	35;	Indels	36;	Gaps
AC	Q918Q3;	Matches	35;	Conservative	18;	Mismatches	55;	Indels	36;	Gaps	
QY	118 EYVLTQPERK 127	118	EYVLTQPERK 127	127							
Db	409 AVIRLWESK 418	409	AVIRLWESK 418	418							
RESULT	13	Query Match	10.7%	Score	73;	DB	2;	Length	463;		
Q918Q3	Q918Q3	Best Local Similarity	24.3%	Pred.	No.	17;	Mismatches	35;	Indels	36;	Gaps
AC	Q918Q3;	Matches	35;	Conservative	18;	Mismatches	55;	Indels	36;	Gaps	
QY	118 EYVLTQPERK 127	118	EYVLTQPERK 127	127							
Db	409 AVIRLWESK 418	409	AVIRLWESK 418	418							
RESULT	13	Query Match	10.7%	Score	73;	DB	2;	Length	463;		
Q918Q3	Q918Q3	Best Local Similarity	24.3%	Pred.	No.	17;	Mismatches	35;	Indels	36;	Gaps
AC	Q918Q3;	Matches	35;	Conservative	18;	Mismatches	55;	Indels	36;	Gaps	
QY	118 EYVLTQPERK 127	118	EYVLTQPERK 127	127							
Db	409 AVIRLWESK 418	409	AVIRLWESK 418	418							
RESULT	13	Query Match	10.7%	Score	73;	DB	2;	Length	463;		
Q918Q3	Q918Q3	Best Local Similarity	24.3%	Pred.	No.	17;	Mismatches	35;	Indels	36;	Gaps
AC	Q918Q3;	Matches	35;	Conservative	18;	Mismatches	55;	Indels	36;	Gaps	
QY	118 EYVLTQPERK 127	118	EYVLTQPERK 127	127							
Db	409 AVIRLWESK 418	409	AVIRLWESK 418	418							
RESULT	13	Query Match	10.7%	Score	73;	DB	2;	Length	463;		
Q918Q3	Q918Q3	Best Local Similarity	24.3%	Pred.	No.	17;	Mismatches	35;	Indels	36;	Gaps
AC	Q918Q3;	Matches	35;	Conservative	18;	Mismatches	55;	Indels	36;	Gaps	
QY	118 EYVLTQPERK 127	118	EYVLTQPERK 127	127							
Db	409 AVIRLWESK 418	409	AVIRLWESK 418	418							
RESULT	13	Query Match	10.7%	Score	73;	DB	2;	Length	463;		
Q918Q3	Q918Q3	Best Local Similarity	24.3%	Pred.	No.	17;	Mismatches	35;	Indels	36;	Gaps
AC	Q918Q3;	Matches	35;	Conservative	18;	Mismatches	55;	Indels	36;	Gaps	
QY	118 EYVLTQPERK 127	118	EYVLTQPERK 127	127							
Db	409 AVIRLWESK 418	409	AVIRLWESK 418	418							
RESULT	13	Query Match	10.7%	Score	73;	DB	2;	Length	463;		
Q918Q3	Q918Q3	Best Local Similarity	24.3%	Pred.	No.	17;	Mismatches	35;	Indels	36;	Gaps
AC	Q918Q3;	Matches	35;	Conservative	18;	Mismatches	55;	Indels	36;	Gaps	
QY	118 EYVLTQPERK 127	118	EYVLTQPERK 127	127							
Db	409 AVIRLWESK 418	409	AVIRLWESK 418	418							
RESULT	13	Query Match	10.7%	Score	73;	DB	2;	Length	463;		
Q918Q3	Q918Q3	Best Local Similarity	24.3%	Pred.	No.	17;	Mismatches	35;	Indels	36;	Gaps
AC	Q918Q3;	Matches	35;	Conservative	18;	Mismatches	55;	Indels	36;	Gaps	
QY	118 EYVLTQPERK 127	118	EYVLTQPERK 127	127							
Db	409 AVIRLWESK 418	409	AVIRLWESK 418	418							
RESULT	13	Query Match	10.7%	Score	73;	DB	2;	Length	463;		
Q918Q3	Q918Q3	Best Local Similarity	24.3%	Pred.	No.	17;	Mismatches	35;	Indels	36;	Gaps
AC	Q918Q3;	Matches	35;	Conservative	18;	Mismatches	55;	Indels	36;	Gaps	
QY	118 EYVLTQPERK 127	118	EYVLTQPERK 127	127							
Db	409 AVIRLWESK 418	409	AVIRLWESK 418	418							
RESULT	13	Query Match	10.7%	Score	73;	DB	2;	Length	463;		
Q918Q3	Q918Q3	Best Local Similarity	24.3%	Pred.	No.	17;	Mismatches	35;	Indels	36;	Gaps
AC	Q918Q3;	Matches	35;	Conservative	18;	Mismatches	55;	Indels	36;	Gaps	
QY	118 EYVLTQPERK 127	118	EYVLTQPERK 127	127							
Db	409 AVIRLWESK 418	409	AVIRLWESK 418	418							
RESULT	13	Query Match	10.7%	Score	73;	DB	2;	Length	463;		
Q918Q3	Q918Q3	Best Local Similarity	24.3%	Pred.	No.	17;	Mismatches	35;	Indels	36;	Gaps
AC	Q918Q3;	Matches	35;	Conservative	18;	Mismatches	55;	Indels	36;	Gaps	
QY	118 EYVLTQPERK 127	118	EYVLTQPERK 127	127							
Db	409 AVIRLWESK 418	409	AVIRLWESK 418	418							
RESULT	13	Query Match	10.7%	Score	73;	DB	2;	Length	463;		
Q918Q3	Q918Q3	Best Local Similarity	24.3%	Pred.	No.	17;	Mismatches	35;	Indels	36;	Gaps
AC	Q918Q3;	Matches	35;	Conservative	18;	Mismatches	55;	Indels	36;	Gaps	
QY	118 EYVLTQPERK 127	118	EYVLTQPERK 127	127							
Db	409 AVIRLWESK 418	409	AVIRLWESK 418	418							
RESULT	13	Query Match	10.7%	Score	73;	DB	2;	Length	463;		
Q918Q3	Q918Q3	Best Local Similarity	24.3%	Pred.	No.	17;	Mismatches	35;	Indels	36;	Gaps
AC	Q918Q3;	Matches	35;	Conservative	18;	Mismatches	55;	Indels	36;	Gaps	
QY	118 EYVLTQPERK 127	118	EYVLTQPERK 127	127							
Db	409 AVIRLWESK 418	409	AVIRLWESK 418	418							
RESULT	13	Query Match	10.7%	Score	73;	DB	2;	Length	463;		
Q918Q3	Q918Q3	Best Local Similarity	24.3%	Pred.	No.	17;	Mismatches	35;	Indels	36;	Gaps
AC	Q918Q3;	Matches	35;	Conservative	18;	Mismatches	55;	Indels	36;	Gaps	
QY	118 EYVLTQPERK 127	118	EYVLTQPERK 127	127							
Db	409 AVIRLWESK 418	409	AVIRLWESK 418	418							
RESULT	13	Query Match	10.7%	Score	73;	DB	2;	Length	463;		
Q918Q3	Q918Q3	Best Local Similarity	24.3%	Pred.	No.	17;	Mismatches	35;	Indels	36;	Gaps
AC	Q918Q3;	Matches	35;	Conservative	18;	Mismatches	55;	Indels	36;	Gaps	
QY	118 EYVLTQPERK 127	118	EYVLTQPERK 127	127							
Db	409 AVIRLWESK 418	409	AVIRLWESK 418	418							
RESULT	13	Query Match	10.7%	Score	73;	DB	2;	Length	463;		
Q918Q3	Q918Q3	Best Local Similarity	24.3%	Pred.	No.	17;	Mismatches	35;	Indels	36;	Gaps
AC	Q918Q3;	Matches	35;	Conservative	18;	Mismatches	55;	Indels	36;	Gaps	
QY	118 EYVLTQPERK 127	118	EYVLTQPERK 127	127							
Db	409 AVIRLWESK 418	409	AVIRLWESK 418	418							
RESULT	13	Query Match	10.7%	Score	73;	DB	2;	Length	463;		

OS Pseudomonas stutzeri (Pseudomonas Perfectomarina)  
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
 OC Pseudomonas;  
 NCBI\_TaxID=316;

[1] RN

SEQUENCE FROM N.A.

RP STRAIN=KC;

RC Lewis T.A., Cortese M.S., Sebat J.L., Green T.L., Crawford R.L.;  
 "Identification of a Region of the Pseudomonas stutzeri strain KC  
 Chromosome Containing Genes for the Biosynthesis of Pyrrole-2,6-  
 bis(thiocarboxylic acid), the Agent of Carbon Tetrachloride  
 Dechlorination Produced by this Organism";  
 Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.

RT EMBL: AF106567; AA/F33139; 1;

DR Interpro: IPR015152; Acyl-CoA\_dh.

PFam: PF0041; Acyl-CoA\_dh; 1.

DR PROSITE: PS0073; Acyl\_CoA\_DH\_2; 1.

DR PROSITE: PS0073; Acyl\_CoA\_DH\_2; 1; UNKNOWN\_1.

SEQUENCE 513 AA; 57432 MW; 4f6d6e16f30d5c34 CRC64;

QY 6 SVRLGVP-----PGRMLQRGIGI--YEDEREWTWTVVRF-NP 41

Db 878 SVPYGVLPRNMRARVVGGRGRCPDLPVGEWLIGGPVANGYRGDPGRT---AERFVND 933

QY 42 SRREVARASQGSYEPSEITHLWQDAH 69

Db 934 DGERWYRSDDLRAYRPGVIEFLGRADH 961

RESULT 14

ID Q9EP5 PRELIMINARY; PRT: 2229 AA.

AC Q9EP5; PRELIMINARY; PRT: 2229 AA.

DT 01-MAR-2001 (TREMBLrel. 16, created)

DT 01-JUN-2001 (TREMBLrel. 17, last sequence update)

DT 01-JUN-2001 (TREMBLrel. 17, last annotation update)

DE POTENTIAL IRON-SULFUR BINDING OXIDOREDUCTASE.

GN SC1B5\_11C.

OS Streptomyces coelicolor.

OC Bacteria; Firmicutes; Actinobacteria; Actinomycetidae;

OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.

NCBI\_TaxID=1902;

RN [1]

SEQUENCE FROM N.A.

RC STRAIN=A3(2);

RA Harris D.; Taylor K.;  
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.

RN [2]

SEQUENCE FROM N.A.

RC STRAIN=A3(2);

RA Parkhill J.; Barrell B.G.; Rajandream M.A.;  
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.

RN [3]

SEQUENCE FROM N.A.

RC STRAIN=A3(2);

RA Redenbach M.; Kieser H.M.; Dempaite D.; Eichner A.; Cullum J.;  
 RA Kinashi H.; Hopwood D.A.;  
 RT "A set of ordered cosmids and a detailed genetic and physical map for  
 RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";  
 RL Mol. Microbiol. 21:77-96(1996).

DR EMBL: AL02351; CAA18985; 1; -.

DR RSPB; 1REFS;

DR Interpro: IPR000025; NAD binding.

DR Interpro: IPR001281; Rieske.

DR PRINTS: P00162; RIESKE.

DR PROSITE: PS00199; RIESKE; 1.

DR PROSITE: PS00200; RIESKE; 2; UNKNOWN\_1.

SEQUENCE 513 AA; 54931 MW; 0f098727cc9f4af CRC64;

QY 11 VPPGRWLRQPG---IYEDEEGR-----TWTWVYRFNSRRIWARASQGSRY 55

Db 436 LPPGEGAVVRAAGGLAVYRDEEGALHAWSPRCTHGLCLVDENAAERAWECPCHSRSF 493

Search completed: November 29, 2001, 04:07:21

us-09-526-329-39.rspt

Page 7

Fri Nov 30 11:03:13 2001

Job time: 222 sec

